THE UNIVERSITY
of EDINBURGH

Epistemic Externalism and the Structure of Justification

Matthew Jope

Doctoral Thesis

The University of Edinburgh

Supervisors:

Dr. Aidan McGlynn
Prof. Duncan Pritchard
Dr. Christoph Kelp

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Matthew Jope	Date

Abstract

This project is concerned with the attempt to diagnose certain types of deductive inferences as exhibiting failure of transmission of justification. The canonical example of alleged transmission failure is G. E. Moore's infamous 'proof' of the external world, in which Moore reasoned *here is a hand, therefore the external world exists*. If the transmission failure diagnosis is correct, then this inference is incapable of providing a route to learning of its conclusion on the grounds that it is only if one already has some justification to believe its conclusion that one can acquire justification for its premise. The thesis presents two novel problems for the transmission failure diagnosis. Firstly, this diagnosis is only appealing if it is able to safeguard epistemic closure, but it is unclear whether it is able to do this once we distinguish between propositional and doxastic justification. Secondly, a dilemma is presented for entitlement theory, the traditional framework for understanding how there could be justification for believing the conclusions at issue. The dilemma arises once we consider the question of what the apposite degree of confidence is that is licenced by an entitlement, the upshot of which is that either entitlement theory is unmotivated, or it leads to irrational doxastic attitudes. In light of these problems, an alternative framework for understanding the phenomenon of transmission failure is presented. This framework rejects a common assumption in the literature that we can only accommodate the phenomenon by appealing to internalist notions of justification. The resulting externalist framework is defended against two potentially devastating objections. Finally, implications for how these views regarding the structure of justification bear on issues in the epistemology of testimony are considered, and it is argued that a certain family of views in the literature on testimony are committed to certain problematic views regarding the structure of justification.

Acknowledgements

I would firstly like to thank my supervisors, Dr. Aidan McGlynn, Professor Duncan Pritchard and Dr. Christoph Kelp for their immeasurably valuable support and guidance throughout my research. I would like to thank the University of Edinburgh for providing the ideal environment for writing the thesis, the *Diaphora Innovative Training Network* for providing financial support and various research activities. Thanks also to the LOGOS Research Group in Analytic Philosophy at the University of Barcelona and the Department of Philosophy at Rutgers University for being such wonderful hosts during my academic secondments.

I would also like to thank Sven Rosenkranz, Michel Croce, Mona Simion, Giada Frantantonio, Crispin Wright, Martin Smith, Ernest Sosa, Matt McGrath, Adam Carter, Dario Mortini, Martin Miragoli, Emma Gordon, Timothy Williamson, Peter Graham, Ram Neta, Sanford Goldberg, Michele Palmira, Guido Tana, Chiara Panizza, Moritz Baron, Ásgeir Berg Matthíasson, Tricia Magalotti, Slawa Loew, Giada Margiotto, Vincent Grandjean-Perrenoud-Contesse, Matheus Valente, Lisa Vogt, Ali Abasnezhad, Jonathan Egeland Harouny, Jonathan Dittrich, Madeleine Hyde, Genoveva Martí, Kathrin Glüer-Pagin, Hannes Leitgeb, , Elisabeth Pacherie, Indrek Reiland, Manuel García-Carpintero, Sònia Roca Royes, Fabrice Correia, Philip Ebert, James Brown, Sophie Keeling, Michela Massimi, Anna Ortin Nadal, Norbert Gratzl, Anna-Maria A. Eder, Anandi Hattiangadi, Peter Pagin, Carl Hoefer, Jesper Kallestrup, Isidora Stojanovic, Peter Sullivan, Åsa Wikforss, Brent Madison, Guy Axtell, Lani Watson, Felipe Morales, Lex Lai, Ricard Guerrero, Doina Bird, Neil McDonnell, Modesto Alonso, Sarah Wright, Tony Lau, Christopher Willard-Kyle, Carolina Flores, Danny Forman, Kegan Shaw, Peter Brössel, Wes Skolits, David Russon, Katherine Palmer, Joey Pollock, Andrés Soria Ruiz, Andy Mueller, Fausto Carcassi, Lilith Newton, Melissa McLaughlin, William McNeil, Katie Keltie, Joe Morrison, Angharad Williams, Stefano D'Orilia, Alex Licht, Phillip Bruner, Bob Jope, Carmel Jope, Jack Kirkham and Rory Kirkham. Thanks also to audiences at conferences, workshops and colloquia at The University of Stirling, the University of Edinburgh, The University of Glasgow, Northwestern University, Durham University, The University of Barcelona, The Jean-Nicod Institute, The University of Wisconsin, The University of Amsterdam and The Bled Epistemology Conference. Thank you to anonymous referees at *The Journal of Philosophy*, *Synthese*, *Erkenntnis* and Routledge. Finally, thank you to Anne & Mick Jope for their unconditional support throughout my academic journey, without which I would surely not have gotten this far.

Some of the materials in this thesis have either been published or are due to be published elsewhere. In particular, parts of chapter 1 are based on 'Intuitive Closure, Transmission Failure and Doxastic Justification' in Jope, M. and Pritchard, D. (eds.), *New Perspectives on Epistemic Closure* (forthcoming[a]). The first half of chapter 4. is based on 'On the Alleged Instability of Externalist Anti-Skepticism,' *The Journal of Philosophy,* (forthcoming[b]). The second half of chapter 4. is based on 'Closure, Credence and Rationality: a Problem for Non-Belief Hinge Epistemology,' *Synthese* (2019). And chapter 5. is based on 'The Symmetry Problem for Testimonial Conservatism' *Synthese*, (forthcoming[c]).

Table of Contents

It is not really difficult to construct a series of inferences, each dependent upon its predecessor and each simple in itself. If, after doing so, one simply knocks out all the central inferences and presents one's audience with the starting-point and the conclusion, one may produce a startling, though perhaps a meretricious, effect.

— Arthur Conan Doyle, “The Adventure of the Dancing Men”, 1903

Introduction

Deductive reasoning presents us with a way of extending our knowledge. By reasoning deductively from known premises, we guarantee knowledge of our conclusions. Or at least, so countless arguments throughout the history of philosophy have assumed. Consider the basic sceptical argument which runs *you cannot know that there is a reality external your mind, therefore virtually all of your beliefs about the world around you fall short of knowledge.* This argument assumes that knowledge that reality is mind-independent is necessary in order to know ordinary empirical truths. The guiding thought behind sceptical arguments like this is that knowledge is *closed* under known deductive inference, which is to say that if one knows that P and that P entails Q, then one cannot fail to know that Q. Classic refutations of scepticism make the same assumptions about knowledge closure too. Consider G. E. Moore's infamous 'proof' of the external world, in which he reasons that *here is one hand, and here is another, therefore the external world exists*. This inference again assumes that if one knows the premise of a deductive inference, one must know its conclusion. When confronted with the Moorean proof, many find it to be less than convincing as a response to scepticism. The question then raised is *where exactly does Moore go wrong*? One answer we can give is that both Moore and the sceptic rely on this principle of epistemic closure and that, notwithstanding its intuitive plausibility, we might think that there are independent philosophical grounds for its rejection. Another option is to argue that what fails is not *per se* epistemic closure but rather a related but logically stronger principle, namely epistemic transmission. Epistemic transmission is the thought that knowledge—or some other epistemic property like justification—*transmits* across inference from premise to conclusion. The difference between closure and transmission is subtle but the key point is that closure might hold while transmission fails in cases where the knowledge or justification for the premise is conditional on one already knowing the conclusion. If, in order to know that, say, *here is an X* on the basis of perception, I first need to know that there is an external world, then the inference *here is an X, therefore there is an external world* will fail to be transmissive. When justification for one proposition depends on prior justification for another proposition in this way, the justificatory architecture is *conservative*.

The conservative diagnosis of what is wrong with the Moorean inference is most closely associated with the philosophy of Crispin Wright. One question that conservatives such as Wright need to consider is where knowledge or justification of the Moorean conclusion might come from if not via transmission from its premise. After all, if one cannot acquire

knowledge of the existence of the external world by means of inferring it from a premise such as Moore's, it looks very difficult to see on what other basis one could know it. The answer that conservatives tend to favour is that the conclusion of Moore's inference has a special kind of status such that we have an epistemic right—an 'entitlement'—to accept it despite our lacking any evidence for it. While entitlement theory represents an interesting response to the sceptical challenge that promises to solve several important questions at once, there is one problem that it faces which I believe has been overlooked in the literature. This relates to the question of how confident one ought to be in a problem on the basis of an entitlement. As I will argue, entitlement theory faces something of a dilemma in response to the confidence question: either it is unmotivated, or it is incoherent.

Wright's account of Entitlement theory is an avowedly internalist picture of warrant. It is often suggested that one of the principal advantages of this sort of internalist approach is that it has the unique capability of offering a viable account of transmission failure. However, with the above dilemma for entitlement theory now on the table, it is worth reconsidering whether we need to engage with internalism in order to account for transmission failure or whether an externalist approach might be viable. A complaint one sometimes comes across in these debates, which we will address in chapter 3., is that externalist accounts of knowledge and justification have a hard time in accommodating the phenomenon of transmission failure. In response to this, I take up the challenge of showing how the externalist does have the tools needed to capture intuitions of transmission failure in at least a decent range of cases. However, the view I advance explicitly rejects the assumption that any plausible theory of transmission failure must count the Moorean inference as non-transmissive. This naturally raises the question of what to say about the initial intuition that the Moorean inference exhibits some kind of circularity or defect. I respond by arguing that the externalist ought to hold that this inference exhibits mere *negative* epistemic dependence and that, while not a bad thing in itself, can have bad epistemic consequences for one who entertains sceptical doubts that one has no reason to entertain.

The resulting view is thus a kind of externalist, neo-Moorean anti-scepticism, which is able to capture our intuitions in the most clear-cut cases of transmission failure. The view will no doubt be unsatisfying to those with prior reservations about externalist theories of knowledge and justification. But notwithstanding such prior reservations, does the account at least offer a coherent response to the sceptical challenge? Wright offers a novel and potentially decisive objection to this type of externalist response to scepticism, arguing that they are ultimately unstable. Curiously, Wright's argument has received relatively little attention in the

literature, despite its potentially devastating consequences for externalism. In chapter 4. I defend externalism against Wright's challenge by showing that it trades on some internalist conceptions about evidence that the externalist would do well to reject. I then go on to defend neo-Moorean accounts in general against an argument from Duncan Pritchard which holds that there is no coherent concept of rationality that would allow for a rational basis for belief in the Moorean conclusion. I show that Pritchard's argument appeals to a principle about rational grounding that we have good reason to reject.

One final question that this thesis is concerned with is how these issues regarding the structure of justification compare across different epistemic domains. On the one hand, it is clear that conservatives intend their account to apply to domains such as testimony and memory, yet on the other hand these examples come up less often. Curiously however, a seemingly separate literature regarding the epistemic status of testimony as a source of knowledge and justification seems to ask more less the same questions regarding the structure of justification. I argue that testimonial conservatives either need to establish important theoretical differences between testimony and perception or they need to be prepared to commit to perceptual conservatism. Testimonial conservatives seem to recognise the potential problem here, wanting to distance themselves from conservatism about perception. However, the arguments they offer either do not establish theoretical differences between perception and testimony or the differences they do establish are not theoretically significant. In light of this, I conclude that testimonial conservatives ought to be prepared to embrace perceptual conservatism.

1.

1. Moore's Argument, Closure and Transmission.

This chapter introduces the distinction between epistemic closure and epistemic transmission, and argues that the attempt to diagnose what is wrong with certain types of inferences in terms of transmission failure faces some serious problems. I begin the discussion by introducing Moore's proof of the external world as an example of the kind of problematic inference at issue in the debate. I discuss two potential strategies for diagnosing the issue namely the rejection of epistemic closure and the rejection of epistemic transmission. I first discuss the closure-rejection strategy, rehearsing the arguments in favour and a variety of well-known problems it faces, noting that the overwhelming intuitive plausibility of epistemic closure in general means this strategy amounts to a kind of epistemological revisionism. I then move on to discuss the transmission failure strategy, and the related conservative account of the structure of justification, one of the principal advantages of which is that it promises to diagnose the issue while safeguarding closure, thereby avoiding any revisionism. I then go on to question the viability of this advantage by arguing that the distinction between transmission and closure is often made by appeal to an outdated formulation of closure that we have independent reasons to want to reject. Once we specify a more intuitive closure principle it becomes much harder to see how conservative accounts of the structure of justification could, in the relevant cases, reject transmission while nonetheless safeguarding closure.

1.1. The Moorean Inference

Consider the following kind of inference due to G. E. Moore. Moore holds his hands up in front of himself and declares 'Here is a hand. If this is a hand, then the external world exists. Therefore, the external world exists.' There is a general sense that this kind of argument is in some sense defective. And yet, it is not immediately obvious why. Moore took this argument to be a perfectly good proof of the existence of an external, mind-independent world in the sense that the inference is deductive, non-circular and the premise is something he knows. Whatever might be wrong with the argument, it is at least a perfectly valid piece of deductive reasoning and one whose premise we would ordinarily take to be known. A challenge, therefore, is to explain the source of the intuition that there is *something* epistemically objectionable about Moore's 'proof.'

A possible answer is that Moore's argument is one instance of a broader class of inferences that exhibit failure of *epistemic closure*. The principle of epistemic closure is roughly the idea that knowledge is preserved, or 'closed', under known entailment. Somewhat

more formally, closure says that if one knows that P, knows that P entails Q, then one knows that Q.[1] Insofar as the Moorean inference seeks to establish not merely that the inference is valid but that its conclusion is known, it does seem to rely on the principle of epistemic closure. While closure seems intuitively very compelling as a general principle, there are some who argue that it fails in certain instances, of which the Moorean inference is one example. If correct, this analysis would present us with a story to tell about what is wrong with the Moorean inference. Although Moore is correct that he knows he has hands, he fails to know what is entailed by this.

Rejecting closure presents us with a story to tell regarding the source of the intuition that Moore's 'proof' is epistemically objectionable in some way. In so doing, at the same time it presents us with a possible strategy for dealing with a certain brand of sceptical argument that is a kind of inversion of the Moorean argument. The sceptical counterpart of the Moorean argument grants the conditional *if one knows that one has hands, then one knows that the external world exists*, while denying that one can know that the external world exists, thereby concluding that one cannot know that one has hands. And since the argument form works, not merely for the proposition that there are hands, but for any ordinary empirical proposition, the radical sceptical conclusion is that we cannot have any ordinary empirical knowledge. We will discuss this closure-based sceptical argument in more detail in chapter 4. but for now, it is helpful just to note that not only would rejecting closure allow us to diagnose what is wrong with inferences such as the Moorean one, but it would also provide us with a possible strategy for responding to a particularly threatening form of scepticism. Both the sceptical argument and the Moorean argument seem to depend on the principle of epistemic closure and thus the rejection of this principle represents one possible strategy for blocking the route to knowledge of their respective conclusions.

We will examine the arguments for abandoning closure shortly, as well as some more examples of alleged cases of closure failure, but first I want to draw attention to the fact that an obvious problem here is that epistemic closure is in general a highly plausible idea. That knowledge is closed across known entailment has an almost 'axiomatic' status in epistemology for many epistemologists.[2] What this means is that any theory that seeks to upend such a strong

[1] As we will see later on, this formulation of closure is overly simplistic and will require refinement. Nonetheless, it will suffice for now as a rough statement of the general idea of closure.

[2] See for example Stewart Cohen (1999, 68) and also Richard Feldman, who takes the rejection of closure to be "one of the least plausible ideas" in contemporary philosophy (1995, 487).

intuition is theoretically revisionary and as such is going to need extremely compelling, independently motivated arguments in its favour.

A second possible way to diagnose what is wrong with an argument such as Moore's is by appeal to failure of *epistemic transmission*. Where closure tells us that knowing that P and that P entails Q is enough to guarantee knowledge that Q, it is silent on where the knowledge of the conclusion comes from.[3] Transmission is a logically stronger principle, which tells us not only is knowledge that Q guaranteed conditional on knowledge that P and the entailment, but that inferring Q from P is a way to *acquire* knowledge of Q. Roughly stated, transmission says that if one knows that P, knows that P entails Q, then one knows that Q on the basis of P and of the entailment.[4] Transmission is thus stronger than closure in the sense that it entails but is not entailed by closure: one might satisfy closure but violate transmission in cases where one knows both premise and conclusion but where the knowledge of the conclusion is acquired independently of the inference. While it is hugely controversial that there are cases in which closure fails, it is trivial that there are cases in which transmission fails. Consider for example an argument that is explicitly question-begging: P therefore P. Such an inference could never be a counterexample to closure and yet it looks like an easy case in which transmission fails. Recognition that some proposition P entails itself necessarily cannot enhance my epistemic standing toward P, whatever that standing is to begin with. Tautologies such as *P therefore P* thus *preserve* but cannot *transmit* knowledge, warrant or justification. A further set of non-transmissive inferences are those for which closure fails. As already made clear, it is controversial that there are cases of closure failure, but *if* there are, then those are again trivially cases of transmission failure given that transmission entails closure. However, this point is, once again, of no use to those who wish to advance the transmission failure diagnosis of inferences such as Moore's because one of the principal advantages of that strategy is that it promises to safeguard closure thereby avoiding the epistemic revisionism attached to the closure-failure strategy. What the transmission failure diagnosis needs to show is that in the relevant cases, closure holds while transmission fails. It needs to be shown, for example, that while Moore does indeed know, or have some justification to believe, that the external world exists, he does not have such knowledge or justification on the basis of the entailment to which he appeals. One problem that the transmission failure diagnosis of Moore's argument

[3] As we will see shortly, this distinction becomes harder to draw once we reflect on more recent 'intuitive' formulations of closure principles.

[4] As with closure, we are going to see below that this formulation is overly simplistic and prone to counterexamples. But for now, it should suffice to express the basic idea behind transmission and how it differs from closure.

faces is thus how to spell out in virtue of what it is that Moore can know or have justification to believe that the external world exists if not via recognition of the entailment from known empirical propositions such as that which figures as the premise in Moore's argument. We will see later on in this chapter, and in more detail in a subsequent chapter, what the proposed solution is to this problem on the part of those who advocate the transmission failure strategy. But the main problem that I will shed light on in this chapter is that while it is ordinarily taken to be obvious that transmission is a logically stronger principle than closure and thus that transmission failure could in principle provide us with a way to solve certain problems while retaining closure, in fact this claim rests on an outdated, questionable formulation of closure which there are independent reasons to reject. Once we articulate a more intuitive closure principle in line with how closure tends to be formulated in the recent epistemological literature, it becomes far less clear how closure and transmission can come apart in the way needed by advocates of the transmission failure strategy.

1.2. On Closure Failure

The debate around epistemic closure goes back as far as 1970 when Fred Dretske published his seminal paper, 'Epistemic Operators'. The radical idea that Dretske puts forward is that there are counterexamples to the idea that knowledge is closed under known logical entailment. Closure seems so compelling in part because deduction cannot introduce any new error possibilities, so that if one has done all that it takes, epistemically speaking, to know that P and one then competently deduces Q from P, thereby acquiring a belief that Q on the basis of the competent deduction, there is a very strong intuition that one has thereby done all that it takes, epistemically speaking, to know that Q. Dretske claims that assumptions of various forms of closure principle pervade the epistemological literature (1970, 1011). And yet, he argues that both sceptics and anti-sceptics are mistaken in assuming that knowledge is closed. Contra this piece of philosophical common sense, there are cases in which one can know P, carry out the relevant deduction to Q, without thereby knowing Q.

To get a handle on the considerations that weigh in favour of rejecting closure, consider Dretske's famous zebra example.[5] You go to the zoo one day and stand before the zebra enclosure. Peering in, you see very clearly what look to be a dazzle of zebras and form the corresponding belief that *those animals are zebras*. Ordinarily, we would be willing to grant

[5] This example, now canonical in the literature on epistemic closure, appears for the first time in Dretske (1970, 1016).

that you know that those animals are zebras based on the evidence available to you. Let's suppose that you do know this. However, if those animals are not zebras, then it is deductively true that they are not mules, and if they are not mules then nor can they be mules cleverly disguised in a visually undetectable way so as to fool unwitting zoogoers into mistaking them for zebras. The question is, do you know that those animals are not cleverly disguised mules? Dretske's suggestion is that your evidence is insufficient, as the case is described, for you to know that they are not cleverly disguised mules. After all, your evidence would appear to be mostly just the look of the beasts, perhaps together with some contextual evidence that zebras are often found in zebra enclosures. But since the mule hypothesis is constructed in such a way that it is their possibility of being disguised in a *visually indiscriminable* way, the visual evidence available to you won't be good enough to rule it out as a possibility. In other words, your knowledge that those animals are zebras entails something that you aren't in a position to know, namely that the animals are not mules in disguise. Ergo, closure must be rejected.

Dretske offers up an array of examples that exhibit roughly this same phenomenon. You know that a certain wall is red, but you don't know that it is not in fact white, cleverly illuminated by red light in such a way that it appears to you to be red (1970, 1015). You know that you have hands, but you don't know that you are not a handless brain in a vat (2005, 14). In each case, there is an entailment from some very ordinary, mundane proposition that is easily known to the denial of some sceptical hypothesis[6] which at face value seems the sort of thing that is much harder to know.

The strategy of rejecting closure promises to provide us with a way to diagnose what is wrong with Moore's argument and at the same time to reject the conclusion of the closure-based sceptical argument. But this offer comes at the price of rejecting a principle that seems intuitively, independently compelling. As mentioned above, one reason closure seems so compelling is that knowledge is factive and deductive inference is truth-preserving, meaning that deduction from known premises guarantees the conclusion is true. Having recognised, via deductive reasoning, that there are no possible states of affairs such that P is true, and Q is false, it seems bizarre that one could know that P is true while not even being in a position to know that Q is true. But beyond these very broad intuitions that speak against the rejection of closure, there are some counterintuitive consequences of the closure-denial view as well as

[6] I mean sceptical hypothesis in the loose sense encompassing, not just the familiar 'evil demon' and 'brain-in-a-vat' scenarios, but also mere local error possibilities such as the hypothesis, mentioned above, that a wall that appears to be red is in fact white, cleverly illuminated to give the appearance of being red.

some problems with the sorts of arguments that have standardly been offered in defence of it. One consequence of rejecting closure is that it seems to have some strange consequences for what it is epistemically permissible to believe and assert about one's epistemic circumstances. If closure fails then presumably it can be reasonable to believe, about oneself, that one fails to know the deductive consequences of what one knows. In other words, the rejection of closure seems to license the following kinds of pattern of reasoning:

> I know that there are two lemons on the table in front of me. There is no doubt in my mind that it is true that here in front of me are two delicious-looking ripe lemons. What's more, I know that lemons are mind-independent objects. However, having taken Fred Dretske's epistemology seminar last semester, I also know that I cannot know there is an external, mind-independent world. So even though I know that there are two lemons in front of me, I don't know that there is a mind-independent world. Similarly, although I know that I took Dretske's epistemology seminar last semester, I don't know that the world didn't come into being a few moments ago complete with all the apparent traces and memories of a deeper history.

The affirmations in this passage sound very odd, albeit short of flat-out contradictions. Following Keith DeRose, affirmations such as these have come to be referred to as "abominable conjunctions" (1995, 28). Abominable conjunctions sound extremely bad. Even Dretske accepts they seem "clearly ridiculous" (2005, 17). He assures us that the closure-rejection project is committed merely to the thought that abominable conjunctions can be true, not that they do not sound ridiculous when uttered. Yet it is not clear from anything Dretske says about them why subjects who realise they are in the grip of a failure of closure ought not be in a position to affirm the relevant abominable conjunction. He assures us that there are other kinds of statements that can be true while sounding absurd such as "The refrigerator is empty but has lots of things in it" (*ibid*). Such an absurd-sounding statement can be true in the sense that each conjunct is true relative to a different context. The refrigerator is empty in the everyday context: it has been emptied of groceries. But it has lots of things in it in a slightly stricter context: there are countless air molecules floating around in it. The comparison here between the contextually variant fridge sentence and an abominable conjunction such as *I know I have hands, but I do not know I'm not a handless brain-in-a-vat* makes sense to the extent that we have already bought into a contextualist account of knowledge ascriptions. But

epistemic contextualism is far from uncontentious[7] and many who find it unattractive are hardly going to be persuaded to abandon closure if it means embracing contextualism as well.

A further set of problems for the closure-denial strategy is that the various attempts at offering independently motivated analyses of knowledge on which closure fails seem themselves to fail. Take for example Robert Nozick's analysis of knowledge in terms of sensitivity, according to which if S knows that P then, were P false, S would not have believed that P (1981, 172). To be more precise, Nozick has in mind a modal account of sensitivity according to which S's belief is sensitive to P iff at the nearest ~P world, S doesn't believe that P. On this account my belief that I have hands counts as knowledge (the nearest ~P world is presumably one in which I lost my hands in some sort of accident, and thus I don't believe I have hands). But my belief that I am not a handless brain in a vat does not count as knowledge: the nearest ~P world here just is the brain-in-a-vat world and it is naturally part of the hypothesis of that world that I have all the same beliefs as I do in the actual world. Likewise, for the other examples of purported closure failure. In each case, one's belief in the ordinary, everyday proposition can be sensitive while one's belief in the entailed proposition is not sensitive. This feels like a good result for those seeking a theory of knowledge on which closure fails in such cases. However, one problem with this account is that it also predicts closure failure in cases where intuitively it should not. The following example is due to Saul Kripke (2011, 186). It is well-known that sensitivity accounts of knowledge predict absence of knowledge in so-called 'fake barn' cases. Suppose I am in an environment where, unbeknownst to me, most of what look like barns are in fact fake barns that have been erected to fool unwitting tourists into believing falsely that they are looking at a barn. Driving past what appears to be a barn I look out of my car window and form the corresponding belief that there is a barn in the field. As it happens, I got lucky on this occasion and this is one of the rare real barns. However, my belief is insensitive. Had there been no actual barn in this field, there would have been a fake barn (let us stipulate), and thus I would still have believed there to be a barn there. Thus, I do not know that there is a barn in the field. But now suppose that the barn I see is a red barn and that all of the fake barns are blue. I now have a perfectly sensitive belief that there is a red barn, since had there not been a red barn, this would have been because there would have been one of the blue fake barns. I thus know that there is a red barn. And yet, I do not know what is entailed by this, namely that there is a barn, since this belief remains

[7] For a survey of the arguments for and against contextualism see Robin McKenna's 'Contextualism in Epistemology' (2015).

insensitive to the truth even while my belief that there is a red barn is sensitive. John Hawthorne adapts this example to make more or less the same point regarding Nozick's sensitivity principle. In Hawthorne's example, one believes truly and sensitively that there is a dog in the room based on seeing a dog, however, just behind the dog out of sight is a very convincing fake cat, which, had the dog not been in the room, would have led one to believe that there is a cat in the room. Thus, although one's belief that there is a dog in the room is sensitive, one's belief that there is an animal in the room is not sensitive, since, had there been no animal in the room, one would have still believed that there was.

One way that a sensitivity theorist might respond to this problem is by specifying that the sensitivity principle is to be understood as a basis-relative notion. The thought would be that we need to hold fixed the basis on which the belief is held across possible worlds and that once we do this it becomes much harder to generate these kinds of problems. Suppose, for example, we stipulate that in the Hawthorne case the belief that there is an animal in the room is formed on the basis of competent deduction from a belief that there is a dog in the room. Now the belief that there is an animal does seem to be sensitive relative to this basis: had there been no animal in the room, one would not have continued to form this belief on the same basis, since one would not have believed there was a dog in the room and thus would not have gone on to deduce that there is an animal in the room. One problem with this kind of response is that it suffers from a version of the generality problem.[8] There are a variety of different ways we can spell out the details of the relevant basis and how we choose to do so can lead to very different results. Suppose that we agree that the basis on which one forms the belief that there is an animal in the room is deduction from a perceptually based belief. Holding fixed this basis, the belief that there is an animal in the room once again fails to be sensitive: had there not been an animal in the room, one would have deduced that there is an animal in the room from one's perceptually based belief that there is a cat in the room. That's a bad result for the sensitivity theorist but things are worse still. Consider what the basis for belief in the conclusion of

[8] The generality problem is the worry is that for any basis-relative epistemic condition such as reliability, safety or sensitivity, a spectrum of options are available for determining which is the relevant process used to arrive at the belief and specifying how coarse-grained or fine-grained we individuate the process, method or basis. And the problem is that there may be different ways to answer these questions, each giving us different answers to the question of whether the belief is reliably (safely, sensitively) based, but it is very difficult to give an independently motivated account of which process is the relevant one and how fine-grained we individuate it. The first extensive discussion of this problem as it applies to reliabilism is Conee and Feldman's 'The Generality Problem for Reliabilism' (1998). Williamson accepts that the generality problem arises for his preferred conditions on knowledge, however he denies that it is a problem, it is simply one manifestation of the way in which the concept of knowledge is vague (2009, 308).

Moore's argument is. If we grant that the basis is deduction from the belief that *here is a hand*, then we get the result that the belief is not sensitive. But consider the option of stipulating that the basis is deduction from *knowledge* of the proposition *here is a hand*. Now the belief is sensitive after all. Had the belief in the external world proposition been false, Moore would not have formed that belief on the basis of deduction from knowledge that he has hands, since Moore would not know he has hands. What the sensitivity theorist who wants to reject closure in the Moorean case but not in these more mundane cases needs to do is offer us an independently plausible account of why the basis for belief ought to be understood in their preferred way. Unfortunately, it seems rather difficult to see how one might go about offering an account in a way that is not woefully ad hoc.

Dretske has an alternative analysis of knowledge to Nozick's that tries to account for closure failure without succumbing to the above problems. According to Dretkse's analysis of knowledge, knowledge requires "conclusive reasons" where this means that R is a conclusive reason for P if and only if R would not be true unless P were true (2004, 19). This counterfactual account can again be understood in modal terms such that R is a conclusive reason for P if and only if all the nearest ~P words are ~R worlds. Crucially, it is not necessary that *all* the ~P worlds are ~R worlds but just the nearby ones. It is again easy to see how this account predicts closure failure in the relevant cases. Suppose that one's reasons for believing *that animal is a zebra* are that one is looking at an animal that has the distinctive zebra look. In such a case, one's reasons are indeed conclusive. For if there were no zebra, then one would not have those reasons. Perhaps the enclosure would have been empty or there would have been some other animal inside, but either way one would not have seen anything that had the distinctive zebra look. However, one's belief that *that animal is not a cleverly disguised mule* is not supported by conclusive reasons. After all, if that belief were false, then one *would*, ex hypothesi, have had the experience of seeing an animal with the distinctive zebra look. Thus, on Dretkse's conclusive reasons account of knowledge, closure fails in the zebra case.[9]

The proposition *I am not a brain in a vat* is what Dretske calls a "heavyweight implication" (2004, 16). These are "known implications to what one perceives (P) that one's perceptual reasons for P are powerless to reach" (*ibid*). What Dretske's conclusive reasons account does well is it predicts that a heavyweight implication such as *I am not a brain in a vat* or *that animal is not a zebra* is not supported by conclusive reasons and thus not

[9] Though similar to Nozick's sensitivity principle, Dretke's account is substantially different and, as John Hawthorne points out, avoids at least some of the problems that arise for Nozick's principle (2005, 35).

knowledgeable. This is, of course, not to say that one could never know these propositions but, rather, that one could never know them given the reasons or evidence specified; closer inspection, expert zoological testimony, or DNA sampling could give one conclusive reasons for believing that the animals are not cleverly disguised mules. While Dretske's account promises to give a plausible story about why closure fails in these but not others, Hawthorne again demonstrates that the account makes the wrong predictions in a range of other cases, predicting closure failure where intuitively it should not. Rather than predict that subjects cannot know in all and only cases involving heavyweight implications, it in fact allows for cases where there are conclusive reasons for heavyweight implications and cases where there are no conclusive reasons despite there being no heavyweight implications. To begin with, here is a case of the second kind (2004, 37).

> I eat some salmon for dinner. I am no glutton. I eat a modest quantity and form the belief that I have eaten less than one pound of salmon. I infer that I have eaten less than 14 pounds of salmon. In fact my perceptual system is very reliable indeed. In those nearby possible worlds where I feast on salmon and eat, say, a pound and a half, I do not believe that I have eaten less than a pound... Suppose (and for all I know this is correct) that while it is utterly unlikely that any human being would eat over 14 pounds of salmon, doing this would induce (among other things) severe hallucinations. Indeed, it might even induce the hallucination that one had eaten a rather small quantity of salmon. Thus it might be true that if, God forbid, I had eaten over 14 pounds of salmon, I may have had the eaten-less-than-one-pound-ish visual experiences that I have at the actual world. Thus while it may be true that I would not have had my reasons R for believing that I have eaten less than a pound of salmon unless I had, it is not true that I would not have had my reasons R for believing that I have eaten less than 14 pounds of salmon unless I had. The apparatus (in conjunction with the verdict that I know, in the given case, that I have eaten less than a pound of salmon) delivers the embarrassing result that I know that I have eaten less than a pound of salmon but not that I have eaten less than 14 pounds (even though I believe the latter and have deduced it from the former).

As with the sensitivity account, the conclusive reasons account again overshoots the mark here, predicting that closure fails in a case where intuitively it ought not. To reiterate the point, the ambition of the conclusive reasons account is that it should rule out knowledge in all and only

cases of heavyweight entailments. However, it has the result that in this case knowledge is ruled out on the grounds of inconclusive reasons even though the entailment is not heavyweight. Whatever value the conclusive reasons account stands to gain by predicting absence of knowledge in the case of heavyweight entailments is surely cancelled out by the disvalue of predicting failure in cases like this where there is a strong intuition that knowledge should be possible. This is indeed an embarrassing result. But things are worse still for the conclusive reasons account, as the following argument from Hawthorne makes clear (2005, 35-36). Recall that according to Dretkse's account I lack conclusive reasons for my belief that I am not a brain in a vat. However, suppose I right now have a headache—then I *do* have conclusive reasons for the conjunction *I have a headache and I'm not a brain in a vat*. After all, the nearest world in which this conjunction is false is one in which I do not have a headache, which is presumably also one in which I do not have reasons for believing that I have a headache, and thus don't have my reasons for the conjunction. So the conclusive reasons account has the peculiar feature that it predicts that I can know—because I can have conclusive reasons for—*I have a headache and I'm not a brain in a vat*—even though I cannot know—because I cannot have conclusive reasons for—I am not a brain in a vat. This looks like an extremely counterintuitive result and shows that Dretkse's analysis of knowledge in terms of conclusive reasons is not able to neatly produce the result that knowledge is ruled out in all and only cases of heavyweight implications.

Both Nozick and Dretske attempt to offer independently motivated analyses of knowledge on which closure fails in all and only the target cases. But both accounts fail to live up to this ambition. The conditions on knowledge that Nozick and Dretske propose predict that knowledge is not closed in the familiar cases involving zebras and brains in vats but, surprisingly, get different results in other cases which seem intuitively to have the same features as the proposed non-closure cases. In addition, Dretske's account also predicts that there is knowledge in cases where, by the lights of his account of heavyweight implications, there ought not be knowledge. Perhaps their conditions could be tweaked so as to eventually match all our intuitions about various cases. However, such tweaks would be ad hoc and thus no longer count as independently motivated. Perhaps an independently motivated account of why knowledge-closure fails is possible, but epistemologists are simply not quite there yet. However, attempts to analyse knowledge in terms of necessary and sufficient conditions in ways that capture intuitions in all cases have a notoriously embarrassing reputation in contemporary epistemology. Given the fact that closure principles in general seem extremely plausible and attempts to capture how they might fail in all and only the relevant cases have so

far been unsuccessful, it is worth considering the possibility that the closure principle is in fact sound.

1.3. On Transmission Failure

An interesting response to debates around the failure of closure comes from Crispin Wright who gives an alternative account of what is going wrong with the inferences in question. The key idea behind Wright's approach is that, in the relevant cases, it is not closure that fails but a stronger principle, namely transmission of warrant or justification. Closure merely states that, given knowledge (or justification) of an entailment from P to Q, if P is known (or justified) then so will Q be known (justified). Transmission is stronger than this. There are a number of ways to state transmission but very roughly the thought is that to acquire justification for the premises of an argument and to recognise its validity is *thereby* to acquire justification for the conclusion. Where closure merely tells us that where there is justification for the premise there will be justification for the conclusion, transmission tells us that justification is acquired by the conclusion in virtue of the inference. It is thus possible for transmission to fail but for closure to hold in cases where it is impossible to acquire justification for Q by inferring it from P because having justification for P already requires there to be justification for Q in place. Wright holds that in cases of alleged failure of closure this is precisely what we find: if one has justification for *those animals are zebras* based on their appearances, then one must already have some justification for *those animals are not cleverly disguised mules* in place, and so closure holds. But now it will be false that the justification one has for *those animals are not cleverly disguised mules* is had in virtue of the justification one has for *those animals are zebras*, given that the former need already be in place to enjoy the latter, and so transmission fails.

A small aside. Closure is normally discussed in relation to knowledge, while transmission is discussed in relation to warrant or justification. This is, however, not necessary. We can happily speak about closure of either knowledge or justification, and likewise about transmission of either knowledge or justification. The crucial distinction is between one principle that states that where one knows that P entails Q, a necessary condition on P being in good epistemic standing is for Q to be in good standing also, and another principle that states that Q will inherit the good standing of P given knowledge of the entailment. This distinction can be made in relation to knowledge or to justification or warrant; however, it is just a fact of how these debates have evolved that, with some exceptions, most of the discussion has been

around closure principles for knowledge and transmission principles for warrant or justification. We will return to this point towards the end of the chapter.

Another small aside. Some philosophers in this debate use the term 'justification', others prefer the term 'warrant'. Usage varies from author to author. Wright sometimes uses the term 'warrant' disjunctively, referring to either evidential justification or non-evidential entitlement[10] (2004, 204). But at other times, Wright seems to use the term 'warrant' as a neutral placeholder for any epistemic condition including knowledge (2008, 509). Alvin Plantinga, on the other hand, uses the term 'warrant' to pick out an externalist condition on belief, which he contrasts to 'justification' which he takes to be internalist (1993). Chris Tucker uses 'warrant' to refer to something that is a reason to believe, while 'justification' is a property of belief (2010). Clearly, these are terms of art and each author is free to specify how they intend to be understood. In what follows, I will mostly use the terms 'warrant' and 'justification' interchangeably, with the exception that where I am explicitly discussing Wright's distinction between warrant-as-justification and warrant-as-entitlement, in which case I will respect his terminology. The reason for my preference is partly that I intend to make use of the well-known distinction between propositional and doxastic justification, in order to talk about different ways in which these two familiar epistemic notions may or may not fail to transmit, but partly also because 'justification' is the more generally familiar term in epistemology and is continuous with the particular theories of epistemic normativity that I will be exploring in subsequent chapters.[11]

Here is how Wright's transmission failure template is intended to work. Consider again the alleged case of closure failure, *here is a hand, there is an external world*. Wright says that what's going on here is not a failure of closure but a failure of transmission of justification across inference. Closure (for justification) says that for any given entailment P→Q, if one has justification for P, then one has justification for Q. Transmission says that the justification one has for P transmits across the inference to Q such that one thereby has justification for Q *in virtue of* the justification one has for P together with the deduction. Transmission is therefore stronger in the sense that it entails but is not entailed by closure. Closure might hold while transmission fails in cases where one has justification for Q, not in virtue of the justification one has for P, but in virtue of some other thing. Trivial cases of transmission failure are

[10] The details of this distinction will be discussed below and again in greater detail in chapter 2.
[11] Particularly in chapter 3., in which I will discuss reliabilist accounts of justification in the context of accommodating transmission failure from within an epistemic externalist framework.

possible. Consider again the tautology *if P then P*. Undoubtedly, if one has justification for P then one has justification for P, so closure holds.[12] But equally certain is the fact that that one cannot acquire a justification—at least not a first-time justification—for P by inferring it from itself, and so transmission fails. What is less obvious is whether there can be non-trivial cases of transmission failure; cases where the entailed proposition is not an explicit part of the entailing premise set. Wright argues that such cases do indeed exist and are precisely those cases of alleged failure of closure.

Consider what justification one might have for a proposition such as *here is a hand*. Presumably some sort of perceptual experience as of a hand. Wright's thought is that whatever justification one has for an empirical proposition such as this it will nonetheless be defeasible. Consider the following inferential structure:

(E) visual experience as of hands
(P) *here is a hand*
(Q) there is an external world to which I have reliable perceptual access

Wright's idea is that E provides justification for P only in a context in which there is already some justification for Q in place (2004, 171). The move from E to a proposition such as P requires, in order for P to be justified, that there is prior justification for Q in place.[13] If this is correct, then clearly transmission will fail in the sense that one will not be able to then *acquire* justification for Q by inferring it from P. The very general sense behind transmission failure is that these inferences fail to provide justification for their conclusions because they exhibit some form of epistemic circularity.

The key idea behind Wright's account of why these inferences are in some sense epistemically circular is that all forms of inquiry necessarily depend upon the good standing of

[12] There will of course be tricky cases in which, simply by the act of inferring, one thereby loses the justification one has for P. For example, suppose that P is I have never performed an inference in my life, and suppose—if we can imagine such a peculiar agent—that one has justification for P. Well then, this will be a case in which simply inferring P from itself will mean one loses the justification one had for P. In claiming that deductively inferring a proposition from itself can never bring one from a justified belief to an unjustified belief, I mean to bracket away apparent counterexamples such as this. In any case, this would not be a failure of closure since there would be no single moment at which one both has justification for P and does not have justification for P.
[13] Though the move from E to P may in certain cases be an inferential move from *the proposition E*, this is not necessary. Wright is happy for his conservative account of the structure of warrant to apply to beliefs formed via some non-inferential process. In such cases we ought to understand E as representing the perceptual experience which provides support for the belief that P rather than a proposition about the perceptual experience from which P is inferred. In mention this here in passing but it will be discussed in more detail in chapter 3.

certain background assumptions.[14] In Wright's terminology these assumptions amount to a commitment that certain *authenticity conditions* obtain.[15] "An authenticity-condition", he writes, "for a given cognitive project is any condition doubt about which would rationally require doubt about the efficacy of the proposed method of executing the project, or about the significance of its result, irrespective of what that result might be" (2014, 215). To doubt an authenticity condition for a particular project may—in cases such as the above where there is an entailment from the P-type proposition to the relevant Q-type proposition—commit one to doubting the specific P-type proposition. But in all cases, it will commit one to doubting the significance of the evidence provided by the E-type proposition for P. For the specific method of forming beliefs based on perception, authenticity conditions include *that there is an external material world, that one's perceptual faculties are reliable*, and so on. To entertain doubts about these propositions would rationally commit one to doubting the evidence delivered by perception.

To hold, as Wright does, that one need, not merely an absence of doubt and absence of grounds for doubt, regarding the authenticity conditions of a given project, but also to *positively accept* that the authenticity conditions obtain and to be justified or warranted in doing so is to treat the justificational architecture *conservatively*. In contrast, to hold the view that it is sufficient that one have no grounds for doubting that the authenticity conditions for a given project obtain is to treat the justificational architecture *liberally*. To stick once again to the Moorean inference, the conservative and liberal thus disagree over whether, in order to have justification to believe the proposition *here are two hands,* one needs prior justification to believe that *there is an external world*. The conservative insists that prior justification is necessary, the liberal insists that absence of grounds or reasons to doubt the existence of the external world is sufficient.

An appealing feature of the transmission failure diagnosis is that it promises to preserve closure while offering an account of why those who had abandoned closure in the first place had been lead to do so: it looked like we had to abandon closure in order to explain an apparent deficiency with certain types of inferences, but this was only because we had failed to recognise the subtle distinction between closure and the stronger principle of transmission. Having made

[14] As we will see below, there is one reading of 'epistemic circularity' which Wright rejects as a general theory of transmission failure, namely the 'background warrant model', according to which the Q-type propositions act as *suppressed premises* in inferences from E to P. A more general sense in which the problematic inferences are epistemically circular is simply that warrant for the Q-type propositions is a precondition of the acquisition of warrant for P-type propositions—this is independent of whether the latter is acquired in some inferential or quasi-inferential process in which the former acts as a suppressed premise.

[15] See, for example, section 11.1 of his article 'On Epistemic Entitlement II' (2014).

this distinction, we can now retain closure while abandoning transmission by embracing a conservative account of the structure of justification. In order to know that *those animals are zebras* one will need to already have some justification in place to believe that *those animals are not cleverly disguised mules*. Prior justification for the latter proposition is part of the presupposed background against which one can acquire justification for the former.

Various stories are possible as to how one could know a proposition such as *those animals are not cleverly disguised* without having carried out some further investigation such as a DNA test or more close visual scrutiny. But let's grant that one could indeed know this on the basis of some kind of background supporting evidence. Closure now holds because one knows both the entailing proposition (on the basis of the visual appearance) and the entailed proposition (on the basis of the background supporting evidence). However, transmission now fails because it is not the case that one has acquired knowledge that *those animals are not cleverly disguised mules* on the basis of the competent deduction, but rather on the basis of the background evidence.

One obvious problem for the transmission failure diagnosis is how to account for where the prior justification for the authenticity conditions comes from. In certain cases, we might well be tempted to agree that there is a huge amount of evidence that one possesses that speaks against the cleverly disguised mule hypothesis. For example, Duncan Pritchard points out that one might reason that "there would be no point to such a deception, that it would be costly, and time-consuming and without bringing any comparable benefit, that it would be easily found out, and then the zoo-owner would be subject to penalties, and so on" (2012, 79). Pritchard distinguishes between discriminating and favouring epistemic support to help with this case (2016, 8-10). Although our evidence may be no help in *discriminating* between the zebra and cleverly-disguised-mule hypotheses, our total evidence, which includes all our background knowledge of how zoos operate, the costs and benefits of deception, and so on, might nonetheless allow us to reasonably *favour* one hypothesis over the other. There are some problems with this approach. Martin Smith has put pressure on Pritchard's strategy for dealing with these so-called 'local' sceptical scenarios (2016). Smith argues that this response to the zebra case will only work against a certain narrow formulation of the sceptical error possibility, while it will be useless against a revised formulation that takes into account one's total evidence. Consider one body of evidence (E^i) *the animals look like zebras* which supports the proposition (P) *the animals are zebras* and an undermining sceptical hypothesis (SH^i) *the animals might be cleverly disguised mules*. It looks like in some sense SH^i, unless it can be

dismissed, undermines the support for P provided by E. Pritchard's strategy is to appeal to a second body of evidence (E^{ii}) *this is a prestigious zoo that would surely not engage in deceptive antics, such a deception would anyway be pointless, costly and time consuming*, which would serve to defeat the sceptical import of SH^{i}, thereby allowing E^{i} to provide justification for P. But a reformulation of SH^{i} which takes E^{ii} into account cannot thereby be dismissed by appeal to E^{ii}. Smith offers up the following alternative sceptical hypothesis: (SH^{ii}) *although zoos are generally trustworthy and reliable and it wouldn't usually be in their interests to deceive the public, the zoo that I'm visiting is an exception to these rules and the animals before me are mules cleverly disguised by the zoo authorities to look like zebras.* Clearly E^{ii} will be ineffective against this reformulated hypothesis which grants all the background evidence one might have regarding how zoos ordinarily operate and proposes that still the current zoo might be the bizarre exception. This suggests that responding to these local sceptical problems by appeal to background favouring evidence is less effective than we might first have thought.

Setting aside these worries about how one might have background evidence that could be appealed to in order to support the conservative diagnosis in the local sceptical cases, an arguably bigger problem is that there seems to be no hope for a similar strategy for the global sceptical scenarios such as the brain-in-a-vat scenario. While the zebra scenario calls into question a very specific body of evidence (the visual appearance of the animals) the more global scenarios target entire domains of evidence (one's perceptual evidence in general). Thus, in the latter type of case one cannot simply help oneself to background supporting evidence in order to justify the relevant type-iii proposition.[16]

The answer that Wright offers to this problem is to propose a type of non-evidential epistemic support that certain background presuppositions can enjoy (2004, 2014). Wright uses the term 'entitlement' to refer to this non-evidential form of warrant, which he contrasts with evidential justification. The general proposal is that we may be entitled to presuppose the truth of certain background assumptions (that we are not brains in vats, and so on) even in the absence of evidential justification for them. An entitlement is a kind of *epistemic right* to presuppose that certain epistemically friendly conditions obtain which are necessary in order to be warranted in forming ordinary beliefs based on evidence. We are entitled to presuppose that we are not brains in vats, that there is an external world, that there are other minds, and so on. We will return to entitlement and its relevance for doxastic justification later on in section

[16] Pritchard has a distinct strategy for dealing with global scepticism which we will address in detail in chapter 4.

1.6. of this chapter, and again in more detail chapter 2. where we will discuss extensively the arguments Wright offers in favour of the rationality of entitlement as well as the range of problems it faces, but for now I want to park these off to one side in order to focus on some more general problems with the transmission failure diagnosis.

The main problem for Wright's transmission failure project that I want to shed light on in the rest of this chapter is the claim that we can keep closure while denying transmission. This claim amounts to one of the central benefits of adopting the transmission failure diagnosis. In order to assess the claim that we can save one principle while rejecting the other, we'll first need to get a bit clearer about the details of the respective principles.

1.4. Intuitive Closure vs. Transmission

Going back to the original Dretske 1970 paper briefly, the general thought behind closure is that epistemic operators such as *knows that, probable that*, etc, penetrate to the known consequences of a proposition. Dretske formulates the principle he has in mind in the following way.

> *Simple Closure*
> If S knows that P, and knows that P entails Q, then S knows that Q[17]

There are a number of things wrong with this initial formulation that are independent of the kinds of problems that Dretske takes himself to have identified for epistemic closure. Crucially, the principle as it stands might fail for the trivial reason that one might have neglected to form any belief in Q. In order to fix this, the principle needs to be refined with the proviso that S believes Q. We will label this formulation *Almost Closure,* in light of the fact that it will need further refinement.

> *Almost Closure*
> If S knows that Q, knows that P entails Q, and believes that Q, then S knows that Q

Almost Closure avoids trivial counterexample due to absence of belief in Q. However, it is still not quite there. One may fail to know that Q if, despite believing that Q, one's belief is not properly based. In order to count as knowledge, one's belief must be based on the right stuff. I

[17] Dretske doesn't actually label this principle 'closure' but rather "Epistemic logic" (1970, 1011).

may know *that oranges are a type of fruit,* know that *if X is a type of fruit, then X is edible*, and even believe that *oranges are edible*, but fail to know the latter because I believe it on the grounds that anything that is orange-coloured is edible. The principle thus needs to be refined again to avoid counterexamples like these in which closure fails but for relatively trivial reasons. The revised formulation will need to specify that the belief that Q is held on the basis of competent deduction from P. However, this introduces an element of diachronicity to the principle, which opens up the possibility for further trivial counterexamples in which S fails to know that Q because, in carrying out the competent deduction, S somehow loses their knowledge that P. In order to safeguard against such counterexamples, we will need to stipulate that S retains their knowledge that P throughout.

In light of these considerations, several authors now formulate closure diachronically, with an explicit emphasis on how closure is meant to capture one way that we may legitimately *extend* our knowledge through inference. Pritchard writes that the reason we ought to find epistemic closure so compelling is that "such principles attempt to codify how one might legitimately extend one's knowledge via competent deduction from what one already knows" (2016, 13). I will borrow the term *Intuitive Closure* from Williamson, to refer to the diachronic formulation of knowledge closure which is meant to capture this idea that inference is a way to extend our knowledge.[18]

> *Intuitive Closure*
> If S knows that P, and S competently deduces Q from P, thereby coming to believe that Q on the basis of the competent deduction while retaining their knowledge that P, then S knows that Q

This formulation avoids the aforementioned problems associated with *Simple Closure* by stipulating that S has competently deduced Q from P. What makes *Intuitive Closure* so intuitive is the thought that deduction is a paradigm way of growing one's knowledge base. Williamson argues from the intuition that "deduction is a way of extending one's knowledge" to the intuitive closure principle (2000, 117). Echoing Williamson in an exchange with Dretske on closure, John Hawthorne argues that "The core idea behind closure is that we can add to what

[18] For various formulations of diachronic closure principles see Williamson (2000, 117), Hawthorne (2005, 29), Silins (2005, 90), Pritchard (2016, 13). The principle stated here most closely resembles Pritchard's, which he simply calls "The Closure Principle" (*ibid*.). I borrow the label from Williamson in order to distinguish it from the rejected formulations above.

we know by performing deductions on what we already know" (2005, 29). Likewise, in trying to formulate a satisfying closure principle, Steven Luper takes us to be trying to capture the intuition that "we can extend our knowledge by recognizing, and accepting thereby, things that follow from something that we know" (Luper, 2016). And in a similar spirit, Duncan Pritchard articulates the force of the intuitiveness of intuitive closure by asking "How could one draw a competent deduction from one's knowledge … without thereby coming to know the deduced conclusion?" (2016, 14). What these remarks highlight is that, in more recent decades, the closure debate is really a debate about extending our knowledge through inference, i.e. about *acquiring* new knowledge through inference.

If we take these remarks at face value and understand closure as a knowledge-extension principle, a problem arises for the transmission failure diagnosis. Recall that the transmission failure diagnosis discussed above boils down to the thought that in certain cases closure holds but transmission fails: one knows Q but the justification one has for Q is independent of the justification one has for P. For this diagnosis to make sense, Wright has to appeal to a synchronic formulation of closure which states merely that "whenever there is warrant for the premises of a valid argument there is warrant for the conclusion too" (2000, 141).[19] In order to get the result that he needs, the version of closure Wright appeals to must allow for cases where one has warrant for the conclusion that is completely independent of the warrant one has for the premise(s). This is clearly not *Intuitive Closure* but is more like the brute synchronic formulations we considered above such as *Simple Closure*. Recall that *Intuitive Closure* stipulates that a subject base their belief in Q on the competent deduction in order to rule out cases in which a subject believes Q but fails to know because their belief has a bad basis. With this restriction, it is hard to see how *Intuitive Closure* will hold in the kind of cases Wright needs in which the justification for the conclusion is completely independent of the justification for the premises and thus independent of the deduction from those premises. This somewhat undermines the benefits of the transmission failure diagnosis that Wright offers because the claim that this strategy can retain closure by rejecting the stronger principle of transmission now comes with a caveat: the kind of closure principle that it can retain is a weaker, unintuitive formulation, which has already been discarded for independent reasons. It is less clear that the transmission failure strategy can reject transmission while retaining *Intuitive Closure*.

In order to appreciate the problem, it will be necessary to have a precise formulation of transmission that we can compare to *Intuitive Closure*. In attempting to articulate a comparative

[19] See also, Wright (2002, 332; 2003, 57).

transmission principle there are a number of ambiguities that we need to address. Firstly, there is the fact that while the above closure principles are, following Dretske and others, stated in terms of knowledge, discussions around transmission and its failure are, following Wright, usually about justification or warrant. In order to assess, therefore, whether closure and transmission can indeed come apart in the way that Wright suggests, we ought to be clear about which epistemic operator or epistemic property we have in mind. We will need therefore to articulate transmission principles for knowledge and for justification.

Secondly, Wright himself distinguishes between a number of different notions of transmission and transmission failure. To begin with, he distinguishes between an inference that is transmissive of first-time justification for its conclusion vs. one that is transmissive of an additional-strengthening justification. These can come apart. Wright supposes that some arguments might be such that they "are essentially apt for warrant enhancement, but not for first-time warrant creation" (2012, 452). The example he uses to illustrate this point is one in which there is a flock of sheep moving about in a small pen and, having counted them, I reach the number twenty-eight and then conclude that the number of sheep is lower than thirty. The thought is that in a context in which I do not already have some justification for this conclusion, the count alone is insufficient to justify me in believing the number is twenty-eight. But in a context in which one has some prior justification to believe the number is lower than thirty, perhaps testimony from the flock's shepherd, then in that context my count now does supply me with additional justification to believe the number is twenty-eight: "the count and the testimonial warrant support each other, and thereby strengthen [one's] overall case for (c)" (2012, 452). So, transmission of first-time justification fails while transmission of additional strengthening justification does not.

In addition to this distinction, another sense in which an inference may exhibit transmission is the notion of being apt to overcome doubt about its conclusion. And yet another notion of transmission is the thought that the very same justification that is had by the premise is had by the conclusion. For example, a visual-perceptual justification which transmits to conclusion providing the same visual-perceptual justification for its conclusion—or not, in the case of transmission failure. There is therefore ambiguity when we introduce the topic of transmission and transmission failure over which kind of transmission principle we have in mind and it is worth being clear about which notion of transmission we are focusing on. In what follows I will begin the discussion of transmission of the first kind, namely transmission in the sense of an inference that is able to provide a *first-time* justification for its conclusion.

One reason for starting here is that in doing so we are in line with most of the literature on transmission and transmission failure.[20]

Finally, an additional complication that arises is the fact that philosophers distinguish between two kinds of justification: propositional and doxastic. Propositional justification refers to what one has justification to believe. Doxastic justification refers to the justificatory status of actual beliefs. Propositional justification for a proposition P does not necessarily entail that one's belief that P is justified. To return to a previous example, I presumably have propositional justification to believe that *oranges are edible*. However, doxastic justification is not guaranteed. I may believe this proposition on a bad epistemic basis (i.e. not on the propositional justification that I in fact have) and therefore fail to have a justified belief despite my having justification to hold this belief. This leaves open the question of whether transmission of justification is best understood in terms of propositional or doxastic justification. Wright himself seems to have something like the notion of propositional warrant in mind when he talks about transmission. For example, in clarifying the concept of transmission he writes "to acquire a warrant *for the premises* of a valid argument and to recognise its validity is to acquire […] a warrant *to accept the conclusion*" (2000, 140-141; emphasis added).[21] Again, the tendency is for the debate to follow Wright in this regard, as Moretti and Piazza confirm: "Most epistemologists investigating transmission and transmission failure broadly identify the epistemic property capable of being transmitted with *propositional* justification" (2018). Nonetheless, there are also interesting questions to be asked about doxastic justification.[22] Not least of all the fact that, to the extent that we are interested in questions concerning knowledge, we ought to be interested, not only in whether an argument could give us justification to believe its conclusion, but also in whether a belief we formed on its basis would itself be justified.

To quickly recap, we now have three kinds of epistemic property about which we want to ask questions: knowledge, propositional justification and doxastic justification. And we have two kinds of epistemic principle that we want to compare: closure and transmission. Given that the question that we want an answer to is whether closure and transmission can come apart for

[20] As indeed Luca Moretti and Tommaso Piazza confirm in their Stanford Encyclopaedia of Philosophy entry on transmission: "Much of the extant literature on epistemic transmission concentrates on examples of transmission of first-time justification" (2018).

[21] And again, for example, drawing the closure-transmission distinction in a subsequent article Wright seems to have in mind something more along the lines of propositional rather than doxastic warrant: "Closure says that whenever there is warrant for the premises of a (known) valid argument, there is warrant for the conclusion too. Transmission says more: roughly, that to acquire a warrant for the premises of a valid argument and to recognise its validity is thereby to acquire—perhaps for the first time—a warrant to accept the conclusion" (2002, 331-332).

[22] For a discussion of transmission of doxastic justification, see Silins (2005).

each of our three kinds of epistemic properties, what we really have are three questions on our hands:

Q1: Can closure and transmission for *propositional justification* come apart?
Q2: Can closure and transmission for *doxastic justification* come apart?
Q3: Can closure and transmission for *knowledge* come apart?

With the questions laid out in this way it becomes easier to appreciate that there are a variety of different ways that the transmission failure diagnosis could be understood, and thus, a variety of different ways in which closure and transmission could potentially come apart. Earlier on we noted some reservations over whether closure and transmission could come apart once we clarify an intuitive closure principle. However, the intuitive closure principle we took was a principle that concerns knowledge, while Wright's characterisations of transmission failure seem to indicate a principle about propositional justification. What we will need in order to substantiate the claim that closure and transmission can come apart are contrasting closure and transmission principles for each of our epistemic properties.

We have already seen that much of the literature on transmission failure has focused on transmission of *propositional* justification or warrant, i.e. on Q1. Wright's conception of transmission seems to be about the acquisition of warrant "to accept" conclusions (2002, 332).[23] Supposing we grant for the sake of argument his account of transmission failure, we might then wonder what follows with respect to Q2 and Q3. Does Wright's account of the structure of justification itself entail corresponding failures of transmission for justified belief and knowledge? One reason for thinking that we want answers to these questions in addition to questions about propositional justification is that most of the initial debates around the possibility of counterexamples to closure focused on knowledge closure. Insofar as the transmission failure diagnosis is supposed to offer an alternative to the closure failure diagnosis, it would be somewhat unsatisfying if it had nothing to say about the knowledge questions. Discussion of the knowledge questions can of course be avoided by embracing a wholesale knowledge scepticism. This is in fact close to views Wright has espoused under the

[23] It is possible to overstate the sense in which Wright's account is an account of propositional justification. His formulations of closure and transmission principles are clearly principles about warrant or justification *to believe* or *to accept* propositions rather than about the epistemic status of agent's doxastic attitudes—as the quoted passages indicate. However, I do not mean to suggest that Wright explicitly intends for his account to apply to propositional warrant or justification rather than doxastic. Instead, all that I am saying here is that the most plausible interpretation of his account is of an account of propositional warrant or justification and that we might do well to inquire how this account extends to doxastic warrant or justification.

label of "The Russellian Retreat" (1991, 88).[24] If we give up the knowledge questions and retreat to questions of justification, then we can simply ignore Q3, but we'll still want an answer to Q2. For this reason, I will focus for the rest of the chapter on answering Q2. In order to do so, it is going to be helpful to clarify the relevant closure and transmission principles. Since those already enjoy much discussion with respect to propositional justification it is going to be easier to start there and then move on to adapt them into principles for doxastic justification.

Before moving on to highlight some problems for the transmission failure diagnosis, I first want to flag a different way that we might construe closure. A possible way to avoid some of the problems that I am going to highlight below might be to embrace an alternative to *Intuitive Closure* in which, instead of strengthening the antecedent in light of the kinds of counterexamples to *Simple Closure* from above, we weaken its consequent. A consequent-weakened closure principle would state that a subject is merely *in a position* to know, or have justification to believe, the relevant Q-proposition.

Closure for Being in a Position to Know (CBPK)
If S knows that P, and knows that P entails Q, then S is in a position to know that Q

This formulation of closure presents an alternative to *Intuitive Closure*, which would seem to avoid the sorts of trivial counterexamples in which a subject fails to know that Q because they fail to form a belief that Q, or because their belief that Q is improperly based, and so on. While there are interesting questions to be asked about consequent-weakened formulations such as CBPK, I will instead focus on the antecedent-strengthened formulations. The reason for this focus is that I want to highlight some problems for the transmission failure diagnosis that arise only once we contrast *having justification to believe* with the *justificatory status* of actual beliefs. And although some of these problems might not arise so easily if we construe closure along the lines of CBPK, to the extent that we are interested in whether and how the actual beliefs that we form on the basis of inferences might be knowledgeable or justified, we ought to reflect on these problems for what I am calling the *intuitive* formulations of closure rather than the *being in a position to know* formulations.

[24] Wright again refers to the idea of the Russellian retreat in (2012, 473) and seems to espouse something like the view again in (2017, 28-29).

1.5. Transmission failure of Propositional Justification

In order to check whether transmission and closure might come apart for propositional justification, we first need to settle on a closure principle for propositional justification. Fortunately, we have the *intuitive closure* principle for knowledge, which can be easily translated into a corresponding principle for propositional justification.

> *P-Closure*
> If S has justification to believe P, and S competently deduces Q from P, while retaining their justification for P, then S has justification to believe Q

There are a couple of differences between *Intuitive closure* and *P-Closure*. Obviously, *Intuitive Closure* mentions knowledge while *P-Closure* mentions justification to believe. But additionally, *P-Closure* weakens the antecedent, dropping the part that required S to have come to believe that Q on the basis of the competent deduction. The reason for this should be clear. It was introduced to the closure principle to avoid trivial counterexamples in which the subject lacks knowledge because they had formed no belief or formed a belief but on the wrong basis. But given that propositional justification for a proposition P does not require that one have formed any belief in P, this part of the principle is not necessary.

Note that *justification to believe* is to be understood here as justification to believe *outright*. If we took *justification to believe* to mean something weaker such as merely having *some reason* to think that P is true, then *P-Closure* would be subject to easy counterexamples. For example, I might have *some* reason to think that a certain playing card is a Queen in virtue of having reason to think it is a face-card. But this clearly gives me no reason to think that it is not a King, despite the obvious entailment. The point here is that we can speak about justification as a gradable notion, where one can enjoy partial justification for P, and justification as an all-or-nothing notion, where one either enjoys justification to believe P outright or one does not. Here and throughout I will use the term 'justification' in the latter sense. The operators, *having some reason to believe that P*, and *having some justification to believe that P*, are clearly not closed, as the cards example nicely illustrates.[25]

I take it that *P-Closure* is highly intuitive. The same reasons that make knowledge closure so intuitively plausible speak in favour of *P-Closure* as well. Since deduction is truth preserving, it seems highly plausible that if one has justification to believe that P is true and

[25] This point is raised by White (2006, 532), McGlynn (2014, 183) and seems to originate in Hempel (1945).

recognises that P cannot be true unless Q is true, then one ought to have justification to believe that Q is true too. Suppose for example that having justification to believe that P is a matter of meeting some threshold where P is sufficiently likely on one's evidence. Meeting the threshold with respect to P and then inferring Q from P would seem to guarantee that one meets the threshold with respect to Q. It is very hard to see how *P-Closure* could fail on most accounts of justification. However, I will not do more to establish the truth of this epistemic closure principle because it will be sufficient to assume it is true in order to answer the question we are interested in, namely whether closure and transmission for propositional justification can come apart such that closure holds and transmission fails.

There are several ways we can formulate our corresponding principle of transmission. Wright sometimes leans towards formulating transmission in terms of *acquiring* a justification for a proposition on the basis of a competent deduction. Here is one formulation he uses:

> (T1) To acquire a warrant for the premises of a valid argument and to recognise its validity is to acquire—perhaps for the first time—a warrant to accept the conclusion[26]

Moretti and Piazza on the other hand formulate transmission, not in terms of justification-acquisition, but instead by stipulating *in virtue of what* it is that one's justification for a proposition obtains. Here is their formulation:

> (T2) If (i) one has justification for P and (ii) one competently deduces Q from P then (iii) one has justification for Q in virtue of the satisfaction of (i) and (ii)[27]

The differences between these two ways of formulating transmission are perhaps not, ultimately, so substantial. T2 is, however, preferable for current purposes because it is more in line with the 'competent deduction' formulation of intuitive closure and will therefore be easier to compare. One thing that we do need to flag is that both of these principles will be prone to one of the kinds of trivial counterexample that we saw above for *Simple Closure*, in which, while carrying out the inference, the subject loses their justification for P. So the transmission principle we need ought to stipulate that the subject retain their justification for P while carrying

[26] (2000, 140-141).
[27] (2018).

out the inference. The transmission principle we will use to compare with *P-Closure* is thus the following:

> *P-Transmission*
> If (i) S has justification to believe P and (ii) S competently deduces Q from P, while retaining their justification for P, then (iii) S has justification to believe Q in virtue of the satisfaction of (i) and (ii)

We are now in a position to compare *P-Closure* and *P-Transmission*. On the face of it, it looks like these could indeed come apart in the way that Wright has in mind. *P-Closure* simply says that if one has propositional justification for P and competently deduces Q from P then one has propositional justification for Q. It doesn't specify where the justification for Q comes from and neither does it make reference to the formation and basing of a belief, so there's no reason to think that it's incompatible with the kinds of cases Wright has in mind in which one satisfies the consequent by virtue of possessing prior, independent justification for Q.

Consider again the obviously question-begging inference *P therefore P*. We noted earlier that we ought to view this inference as trivially non-transmissive. If our closure and transmission principles are to be viable, one thing they ought to get right is that this inference satisfies closure but violates transmission. This is indeed what we find. First assume that there is justification for the premise, so justification for P. Since the premise of this inference is identical to the conclusion, and the antecedent of P-Closure stipulates that, while carrying out the inference, S retains their justification for the premise, it looks like there is no logical space for this inference violating P-Closure. So far so good. But now see that it is equally obvious that this inference will violate P-Transmission. First assume that the antecedent of P-Transmission is satisfied such that S has justification for P. Now it will be impossible to satisfy the consequent, given that P will not be justified in virtue of the inference but in virtue of whatever it was that made it justified prior to carrying out the inference. Thus, our two principles, P-Closure and P-Transmission, look like they are perfectly well able to accommodate the non-transmissivity of the obviously question-begging inference. Notice also that we did not need to appeal to any kind of conservative account of the structure of justification to get this result. This is unsurprising because the whole point about the existence of trivially non-transmissive inferences is that all parties to the debate ought to be able to acknowledge them.

The next thing to do is to check whether our principles can accommodate transmission failure in other kinds of cases as well, such as those that Wright takes to exhibit transmission failure. Recall that the basic idea behind transmission failure in the kinds of cases that Wright has in mind is his conservative account of the structure of warrant. The rough idea is that certain propositions—authenticity conditions—function as *presuppositions* of inquiry. The justification that I have to believe that I have hands depends on there being a corresponding perceptual experience, but this experience alone is not enough to give me justification. After all, I could have the experience and yet be a handless brain in a vat. The evidential support needs to be backed by further justification to presuppose that I am not dreaming, that I am not a brain in a vat, and so on. The notion of presupposition here is a weak one in the sense that it does not require that one form an actual belief. Wright suggests that the correct way to understand our commitment towards these propositions is a kind of implicit trust (2004, 192-193). Nonetheless, conservatism requires that there already be justification in place for these presuppositions. And for this reason, transmission has to fail in all cases where there is an inference from an evidentially justified proposition P to an authenticity condition for the method of inquiring into P based on the kind of evidence in support of P. Q cannot be justified *in virtue of* the justification for P if justification for Q is among the background conditions that make P justified in the first place. For example, Wright thinks that prior background justification for *there is an external world* is necessary in order to acquire justification for *here is a hand*. On the supposition that this account of justification is correct, it follows that we have a case in which *P-Closure* is satisfied but *P-Transmission* is not. One might have justification for *here is a hand* and also for *there is an external world*, and so satisfy the antecedent and consequent of *P-Closure*. But given that justification for *there is an external world* is among the background conditions that make *here is a hand* justified in the first place, it won't be the case that one has justification for *there is external world* in virtue of the justification one has for *here is a hand*.

Pryor disagrees with the conservative diagnosis of the Moorean inference. He argues that a proposition such as *here is a hand* can be *immediately* perceptually justified without the need for prior justification for *there is an external world*. However, it is interesting to note that Pryor does take the conservative diagnosis to apply in other cases, such as the zebra inference. Consider again the zebra inference from earlier.

(P) that animal is a Zebra

(Q) that animal is not a cleverly disguised mule

Pryor and Wright agree that we ought to treat the justificatory architecture of this inference conservatively and thus that it is a paradigm case of transmission failure. With this in mind, it is important to recognise that these labels apply to not across the board but to particular triplets consisting of some kind of ordinary, defeasible evidential justification (E), a proposition (P) supported by E and another proposition (Q) doubt about which would undermine the support for P provided by E. In labelling one view as conservatism and another as liberalism, we need to be clear that what we really mean is that a view counts *this-or-that* triad conservatively or liberally. As Wright points out, there need be no "global" opposition between conservatives and liberals regarding all justificatory triads (2014, 218). The question is rather whether, in any particular case, we should treat the justificatory architecture liberally or conservatively.[28] This relates to another difference between closure and transmission. When we ask whether closure and transmission can come apart such that closure holds and transmission fails, what we are really asking is whether this applies of a particular case. What we want to know is whether, with respect to a *certain inference*, it is possible for closure to hold and transmission to fail. On Wright's account of the structure of basic perceptual propositional justification the answer is yes with respect to the Moorean inference, the brain-in-a-vat inference and the zebra inference. Pryor's 'liberal' account differs from Wright in that he does not think there is transmission failure in the Moorean and brain-in-a-vat inferences, though he is happy to accept conservatism, and thus transmission failure, with respect to the zebra inference.

1.6. Transmission Failure of Doxastic justification

We have seen how Wright's account of transmission failure works for propositional justification such that it has the result that closure for propositional justification and transmission for propositional justification can come apart. We noted that there is some disagreement over which arguments exhibit the phenomenon of transmission failure, but it is interesting that those who defend traditionally *liberal* accounts of justification such as Pryor do countenance transmission failure in certain cases. So the answer to Q1 above seems to be a resounding, yes. What about Q2? As before, we'll need to specify the relevant closure and transmission principles for doxastic justification first before examining whether they may indeed come apart. Again, we can adapt our *Intuitive Closure* principle for knowledge into a corresponding principle for doxastic justification:

[28] In referring to Wright as a *conservative* and Pryor as a *liberal*, what I and, I believe others, most often intend is that these authors treat the specific Moorean inference conservatively or liberally.

D-Closure

If S's belief that P is justified, and S competently deduces Q from P, thereby coming to believe that Q on the basis of the competent deduction, while retaining their justified belief that P, then S's belief that Q is justified

Likewise, we can formulate our transmission principle for doxastic justification by adapting the above transmission principle for propositional justification into a principle for doxastic justification:

D-Transmission

If (i) S's belief that P is justified and (ii) S competently deduces Q from P, thereby coming to believe that Q on the basis of P, while retaining their justified belief that P, then S's belief that Q is justified in virtue of the satisfaction of (i) and (ii)

What then is to be said about whether closure and transmission can come apart for doxastic justification now that we have clarified the relevant principles? The first thing to note is that *D-Closure* contains the strengthened antecedent—*thereby coming to believe that Q on the basis of P*—which was part of *Intuitive Closure,* but which was jettisoned when it came to formulating *P-Closure*; jettisoned because propositional justification, unlike knowledge and doxastic justification, does not require that the subject have actually formed a belief in the relevant proposition. But things are different for doxastic justification since it involves the justificatory status of actual beliefs. Hence, *D*-Closure is more similar to the intuitive closure principle for knowledge than was *P-Closure*. With the strengthened antecedent now in play, it becomes more difficult to see how the conservative framework can get the result that closure and transmission could come apart for doxastic justification in the relevant cases, even on the assumption that they do come apart for propositional justification in those cases.

To appreciate the problem, we will first assume that *D-Closure* holds. This assumption is both reasonable—since *D-Closure* is extremely plausible—and necessary if the transmission failure project is to be vindicated. After all, notice that if *D-Closure* does not hold, then it is hard to make sense of how *P-Closure* could nonetheless hold. In other words, it is hard to make sense of how there could be justification to believe a proposition which could never be justifiably believed (something that would be the case for conclusions of the types of inferences under discussion if closure holds for propositional justification but fails for doxastic justification). Moreover, one of the main benefits of the transmission failure project is that it

promises to avoid a range of problems for the Dretske-style denial of closure for knowledge, but by denying closure for doxastic justification we simply reintroduce many of these same problems but for justified belief rather than knowledge. Furthermore, to suggest that closure holds for propositional justification but fails for doxastic justification is to entertain the thought that there could be justification to believe a proposition that could never be justifiably believed. It is unclear whether this a coherent possibility. It is tempting to think that if there are reasons for thinking that a proposition could in principle never be justifiably believed, then there could never be justification to believe it. All this it to say that we may reasonably assume that *D-Closure* holds for the purposes of assessing the transmission failure diagnosis.

Assuming that *D-Closure* holds, is it possible for *D-Transmission* to fail? The first thing to note is that it will again be trivial that the obviously question-begging inference violates D-Transmission, and that this will be for much the same reason as for P-Transmission. Effectively, it boils down to the fact that if one has a justified belief that P, infers P from itself while retaining one's justified belief that P, then one's belief that P will be justified (so D-Closure holds). However, the belief that P, assuming it is still justified, will be justified in virtue of whatever it was that justified it in the first place, not in virtue of the inference (and so D-Transmission fails). The *assuming it is still justified* is important. One might object to the foregoing reasoning, insisting that once our subject infers P from itself and thereby bases their belief that P on the question-begging reasoning, their belief that P loses its justification. After all, it was justified initially because it was based on, say, good evidence, but if, having inferred the proposition from itself, one now bases one's belief that P on P's self-entailment rather than the evidence for P, this looks like a different basis, so there is no need to assume the belief retains its justified status. If this is correct, then the question-begging inference threatens to violate, not just D-Transmission, but D-Closure, which looks like a bad result. However, if we bear in mind that the antecedent of D-Closure specifies that, while carrying out the inference from premise to conclusion, the subject retain their justified belief in the premise, then D-Closure is not violated in this way. If the case under consideration is one in which, by carrying out the inference, one loses one's justification for the premise, then it will not satisfy the antecedent of D-Closure and thus not violate D-Closure. Ergo, the simple question-begging inference satisfies D-Closure but violates D-Transmission, which is indeed the desired result.

The more interesting question is whether, with these doxastic principles in play, the conservative can maintain that there are non-trivial cases of transmission failure. If so, then the thought has to be something like this. One can satisfy closure by having a justified belief in both the entailing and entailed proposition, but where the latter is justified independently of

any kind of inference from the former. A challenge for such an account then is to spell out what it is in virtue of which a belief in a proposition such as *I am not a brain in a vat* could be justified if not inferentially from a justified belief in a proposition such as *here is a hand*. Perhaps a similar conservative story could be told here to the one told above for propositional justification. Recall that the answer to the corresponding question for propositional justification was that there could be a type of default warrant or justification to presuppose that one is not a brain in a vat. This default epistemic status is in no way a type of evidence but, according to Wright, is to be understood as a kind of *epistemic entitlement* or *epistemic right*. It is our right to trust implicitly that certain sceptical scenarios are false. 'Implicit' in the sense that such trust need involve no belief in, nor indeed any awareness of, the proposition in question. Nonetheless, the entitlement to presuppose that one is not a brain in a vat needs to be in place in the structure of justification in order for one to initially acquire justification for *here is a hand*, and it is for this reason that propositional justification fails to transmit across the inference from *here is a hand* to *I am not a brain in a vat*. If a similar story is going to work for doxastic justification, then it will be necessary to make sense of how it is that a belief in a proposition such as *I am not a brain in a vat* could be justified independently of any deductive inference from a belief in a proposition such as *here is a hand*.

On standard accounts, doxastic justification is explained in terms of propositional justification. The thought is roughly that a belief is justified if and only if it is a belief in a proposition for which there is justification and it is based on that which propositionally justifies its content.[29] So, for example, a belief in a proposition such as *it is raining outside* is justified if and only if (a) there is justification to believe the proposition *it is raining outside* and (b) the belief is based on whatever justifies this proposition (in this example, a perceptual experience as of it raining outside). So, for a belief in a Q-proposition to be (doxastically) justified it needs to meet these twin conditions. We can assume that condition (a) is met given that propositional justification is closed. How about condition (b)? One difficulty here is that the kind of propositional justification that obtains for Q-propositions is not of the ordinary, evidential kind,

[29] The standard account of doxastic justification has come under pressure from John Turri who argues that we ought to run the analysis in the other direction, understanding propositional justification in terms of doxastic (2010). The argument I run in this section works under the assumption of the standard account, but nothing hinges on this. For fans of Turri's account, the relevant question to ask in order to assess whether *D-Closure* holds while *D-Transmission* fails is whether one might possess some means of coming to believe that *q*, independently of inference from *p*, such that, were one to believe that *q* via one of those ways, one's belief would thereby be doxastically justified. It should be clear that on the entitlement theoretical account, the answer here is a resounding *no*. Although one is entitled to trust that *there is an external world*, if entitlement theory and the failure of transmission of justification framework that comes with it is correct, then one does not possess a means of coming to justifiably believe in the existence of the external world.

but rather the default entitlement kind, and it is not immediately obvious that an epistemic right—an entitlement—is the correct sort of thing on which to *base* a belief. As discussed, Wright seems to understand the notion of an entitlement to Q along the lines of a right to *trust* that Q rather than a justification to *believe* that Q. In a very strict sense, then, even closure does not hold for doxastic justification, let alone for transmission.[30] However, the same is true of Wright's account of propositional justification in the sense that evidential justification to believe is not closed.[31] Wright manages to rescue closure only by defending a disjunctive account of justification (or warrant), according to which a proposition can be justified if it is either justified evidentially or justified non-evidentially. Perhaps, then, the more charitable way to a corresponding account of doxastic justification is to see doxastic justification as a disjunctive notion also. Building off of Wright's disjunctive conception of propositional justification, we may say that doxastic justification comes in two varieties: either it is a matter of a belief that P being based on that which justifies P, or it is a matter of implicit trust in P where there is an epistemic right to trust.

With the disjunctive account of doxastic justification now in play, we can try to make sense of what a conservative account of doxastic justification might look like. Closure for doxastic justification could hold if one were to have a doxastically justified belief in the P-proposition and a doxastically justified (entitled) trusting in the corresponding Q-proposition. We could then treat an inference conservatively by insisting that in order to justifiably believe the premise one needs to already have implicitly trusted in the conclusion and to be justified (entitled) in doing so. Finally, it seems like such an account would have the consequence that

[30] This very clearly follows from the view that Wright avows in his first at-length defence of epistemic entitlement (2004). He rehearses this account again a decade later in his second entitlement paper (2014), but, in light of pressure from McGlynn (2014) vis-à-vis the alchemy problem, is forced to go on to concede that there may be transmission after all, which entails that closure for doxastic justification holds after all.

[31] Though this is unambiguously the view Wright advances in his 2004 entitlement article, two caveats are needed. Firstly, although the disjunctive account of warrant entails that *justification to believe* is not closed, the reason for this is one that, at a certain point, Wright seems not entirely wedded to. If indeed the concept of 'belief' is evidentially controlled, then justification to believe is not closed. Wright, however, tells us "I do not myself know whether the notion of belief *is* actually so tightly evidentially controlled" (2004, 176). Rather, he seems to prefer to grant this as an assumption, meaning that the claim that justification-to-believe is not closed is ultimately conditional on this assumption. Secondly, in his subsequent 2014 article, in response to concerns regarding the alchemy and leaching problems, Wright's considered view seems to be that there is closure for justification-to-believe after all, however the resulting notion of belief in Q-type propositions is one whose evidential credentials "depends for its force on our antecedent reason to trust in the truth of those very same propositions, so that the ultimate authority we have for accepting them depends on the rationality of that trust" (2014, 235). I discuss this admission and what it means for Wright's overall strategy in more detail in chapter 2, however I stick with the original account in this chapter for the reasons that it helps us to get a handle on the main thrust of the disjunctive conception of warrant that is central to Wright's view and that, as I will detail in chapter 2, it is unclear how to make sense of Wright's subsequent admission that there is closure for justification-to-believe after all.

the trust that one places in the conclusion is justified (entitled) independently of the justification one has for the inference, which means that on this account of doxastic justification, closure holds, and transmission fails.

The problem with this disjunctive account of doxastic justification is that, while it secures closure in this loose sense of there being doxastically justified attitudes towards both premise and conclusion of an inference, it does not secure the result that *D-Closure* holds. This is because *D-Closure* specifically mentions the formation of a belief—not merely an implicit trusting—which is *based on* an inference. So whether or not one is entitled to implicitly trust that Q, in order for *D-Closure* to be respected it will be necessary that one has a doxastically justified belief in the conclusion of the relevant inferences and that this belief is based on the inference. And this is precisely what the conservative account of the structure of doxastic justification, which we have just tried to develop, is committed to denying. And because we registered above the point that a basic requirement of any transmission failure strategy is that it safeguards closure, this conservative account of the structure of doxastic justification fails to present a viable transmission failure strategy for doxastic justification. Thus, the answer at this point to Q2 seems to be that closure and transmission for doxastic justification do not come apart in the same way that conservatives argue that closure and transmission for propositional justification come apart—the disjunctive account of doxastic justification notwithstanding.

One way that a defender of the disjunctive account might respond here is to point out that *D-Closure* is a peculiar rendering of the closure principle which was only brought in earlier on in this chapter as a kind of interpretation of a doxastic version of the *Intuitive Closure* principle for knowledge. One might be tempted to think that we can make do with rejecting *D-Closure* if we can nonetheless retain some other, looser version of a closure principle for doxastic justification that does not require the formation of an actual belief as *D-Closure* does. Unfortunately, there are reasons to think that the problems here are unavoidable. What the conservative account under consideration can safeguard is a closure principle that says that the kind of justification one must have with respect to Q-propositions will be of the entitlement variety. That is, a mere justified implicit trusting in propositions such as *I am not a brain in a vat*. But the unavoidable issue for any such account is what to say about a subject who has recognised the deductive consequence of their belief that they have hands, thereby forming a belief that they are not a brain in a vat. The conservative account of doxastic justification seems committed to denying that the resulting belief could ever be justified. And insofar as an account is committed to *that*, it is, in very clear sense, committed to denying closure for justified belief. It is hardly a consolation to show that we can retain closure in this disjunctive sense once we

recognise that in reasoning deductively, we make the implicit explicit by teasing out the deductive consequences of our beliefs.

1.7. Concluding Remarks

This chapter began by introducing a certain type of allegedly problematic inferential pattern, exemplified by the Moorean argument. Two broad strategies were surveyed for diagnosing the problem, the rejection of epistemic closure and the rejection of epistemic transmission. It was noted that the rejection of closure is highly theoretically revisionary and arguments that support this strategy face a range of well-known problems. I then went on to discuss a related but distinct strategy, namely the rejection of epistemic transmission. The alleged benefit of the transmission failure strategy is that it promises to safeguard closure while rejecting the logically stronger principle of transmission of justification or warrant. I noted that one important question facing the transmission failure response will be kept for a later chapter, namely the question of where justification for propositions such as *I am not a brain in a vat* comes from if not via closure-style deductions from propositions such as *here is a hand* (§3.). I then went on to question one of the purported benefits of rejecting transmission, namely the thought that by doing so we can safeguard the logically weaker closure principle. I noted that this claim depends on an outdated formulation of closure that few in the literature would nowadays accept. After distinguishing the outdated formulation from the more up-to-date *intuitive* closure principle, I then went on to draw an important distinction between closure and transmission of different kinds of epistemic properties, namely knowledge, propositional justification and doxastic justification. I ended by showing that even on the assumption that transmission fails for propositional justification it is hard to see how, in the relevant cases, transmission could fail for doxastic justification in such a way that manages to safeguard closure. The upshot is that, once we bear in mind that it is the so-called *intuitive* closure principles that we need to safeguard, Wright's conservative account of the structure of justification seems to have a problem in predicting that transmission fails in the relevant cases but not closure. In subsequent chapters I will address questions such as how the Moorean inference relates to the problem of closure-based-scepticism, what advocates of the transmission failure strategy want to say about where justification for the conclusions of the relevant inferences might come from if not from transmission, as well as what options there are for externalists who want to safeguard closure while rejecting the internalist assumptions at the heart of the transmission failure diagnosis.

2.

2. A Dilemma for Entitlement Theory

Consider again the Moorean inference: *here is a hand, therefore the external world exists*. The conservative picture tells us that the reason this inference is non-transmissive is that warrant for its conclusion is part of what gives one warrant for the premise in the first place.[32] Warrant for the conclusion is *epistemically prior* to warrant for the premise, in the sense that one first needs warrant for it if one is to acquire warrant for the premise. Naturally, this raises the question of where warrant for a proposition such as *the external world exists* could come from if not from the kind of deductive reasoning exemplified by Moore's inference. The sceptical conclusion to draw is that it is therefore impossible to acquire warrant for propositions such as *the external world exists*, given that it cannot be acquired via transmission from empirical propositions such as *here is a hand*, and that there seems to be no non-empirical alternative route to doing so. This looks like a devastating sceptical result. However, Wright points out that we are not yet at the conclusion the sceptic needs to draw, namely that the proposition *there is an external world* is unwarranted. It need not follow from the fact that there is no way to *acquire* warrant for this proposition that we do not *possess* a warrant for it.[33] If we could make sense of a type of warrant that does not need to be acquired because it is possessed by default, then the needed sceptical conclusion is blocked. Here is Wright gesturing at what he has in mind:

> Suppose there is a type of rational warrant which one does not have to do any specific evidential work to earn: better, a type of rational warrant whose possession does not require the existence of evidence—in the broadest sense, encompassing both a priori and empirical considerations—for the truth of the warranted proposition. Call it entitlement. If I am entitled to accept P, then my doing so is beyond rational reproach even though I can point to no cognitive accomplishment in my life, whether empirical or a priori, inferential or non-inferential, whose upshot could reasonably be contended to be that I had come to know that P, or had succeeded in getting evidence justifying P.[34]

[32] On the distinction between 'warrant' and 'justification.' As discussed in the previous chapter, I take these to be terms of art that each author is free to define as they deem useful. I prefer the term justification and substituted 'transmission of warrant' with 'transmission of justification,' which suited the purposes of the previous chapter. However, this chapter deals so heavily with Wright's distinction between two types of warrant—evidential justification and non-evidential entitlement—that I revert here to his terminology.
[33] Wright refers to this as a theoretical "lacuna" in the most interesting forms of sceptical argument (2004, 174).
[34] Wright (2004, 174-175).

Entitlement theory thus promises a response to scepticism in the form of a species of warrant, distinct from the familiar conception of evidential justification, that one does not need to do any specific evidential work to earn or acquire. So long as we have no evidence to believe they are false, we may be entitled—warranted by default—to certain of our anti-sceptical commitments. This chapter looks in detail at Wright's account of entitlement, at some of the well-known problems it faces that others in the literature have identified, and then discusses a novel problem for entitlement theory in the form of a dilemma. This dilemma arises in light of Wright's answer to the question of what exactly an entitlement is a warrant to do. Entitlements are warrants to place *trust* in the truth of certain propositions (Wright, 2004, 2012, 2014). Trust is here understood as a species of acceptance that "comes apart from belief in cases where one is warranted in [accepting] that P for reasons that do not bear on the likely truth of P" (2004, 177). As we are going to see, it is not clear how best to understand the type of doxastic state and associated degree of confidence that Wright intends by this notion of propositional trust. If trust based on an entitlement involves a relatively low degree of confidence, this seems to licence irrational doxastic attitudes when combined with belief in the outputs of the kinds of inquiries for which entitlements function as presuppositions. On the other hand, if entitlements are warrants to be very highly confident, then it seems like they are not well supported by the kinds of arguments typically on offer in their defence. Entitlement theory is thus either irrational or unmotivated. Such is the dilemma that I will defend towards the end of this chapter. But before getting to that, we will first explore Wright's account of entitlement, the proposed arguments in favour of the rationality of entitlements to trust in our anti-sceptical commitments, and some of the already well-established arguments against entitlement theory.

2.1. Entitlement to Trust

The entitlement project begins with the reflection that evidential inquiry necessarily involves taking on board a range of presuppositions, which Wright refers to as *authenticity conditions* (2014, 214). In order to form a belief based on evidence, one is committed to taking it that certain epistemically-friendly conditions obtain. Such presuppositions or assumptions are not themselves the outputs of any prior evidential inquiry but, rather, need to be in place first in order for inquiry to begin. These propositions act as 'hinges' for inquiry in the sense that inquiry turns on them and depends on them holding steadfast, much in the same way that a door cannot properly turn without its hinges staying put.[35] To doubt such a proposition would

[35] The hinge metaphor originates in Wittgenstein (1969).

call into doubt a whole range of other propositions, or of the significance of the evidence one has in support of them, for which it functions as a presupposition. "An authenticity-condition for a given cognitive project," Wright explains, "is any condition doubt about which would rationally require doubt about the efficacy of the proposed method of executing the project, or about the significance of its result, irrespective of what that result might be" (2014, 215). Many authenticity conditions will be such that they *could* in principle be investigated, although that investigation would itself be its own project that would again come with its own authenticity conditions. For example, in testing the temperature of a liquid and arriving at a certain result, the question of the reliability of the thermometer is not part of the investigation but is itself take for granted. One could of course perform a test to see whether the thermometer is reliable, but such a project would come with its own presuppositions. Other authenticity conditions will be more general and will be in place for virtually all possible forms of inquiry within a certain domain. So for the domain of empirical inquiry, one is committed to presupposing that one's perceptual faculties are reliable, for example. The general point is that, for all authenticity conditions, we are in some sense committed to the implicit acceptance of these presuppositions when engaging in inquiry, even though such acceptance may not itself be the output of any prior inquiry. Another important (in the context of the current discussion) example is the by now familiar conclusion of Moore's argument *that there is an external world*.

A brief terminological point before continuing. In what follows I will be using the terms 'hinge' and 'authenticity condition' somewhat interchangeably. The latter is Wright's preferred term and perhaps more accurately fits his tripartite account of the structure of justification in which some defeasible evidence supports belief in a quotidian proposition only in a context in which there is warrant to accept that certain epistemically-friendly conditions obtain. The notion of a 'hinge' is more widely used in the literature but carries with it certain ambiguities. For example, the notion is sometimes used to pick out the kind of proposition that plays the role of an authenticity condition such as the conclusion of Moore's inference. At other times, the notion is used to refer to propositions about which one is optimally certain, which includes such things as the *premise* of Moore's inference. Where I will be using the terms 'hinge' and 'authenticity condition' interchangeably, I intend the more restricted notion of a hinge as a background assumption of inquiry such as *that one's perceptual faculties are reliable, that there is an external world,* and so on.

The proposal that inquiry necessarily involves the implicit acceptance of a range of authenticity conditions that are not themselves the output of that inquiry naturally raises the

question of what epistemic status these presuppositions have. Consider that in forming beliefs based on perception, I am committed to the conclusion of Moore's argument:

(Q) There is an external world

In order for the practice of forming beliefs based on perception to be in good standing, I must be warranted in accepting authenticity conditions such as Q. So, a warrant for Q is required, prior to the acquisition of warrant for any propositions that are outputs of the kind of inquiry for which Q functions as a presupposition: propositions such as *here are two hands, there is a cup on the table, it is raining outside*, and so on. The problematic thought is, where could such a warrant for Q come from given that there is no way to investigate the truth of Q that wouldn't already depend on presupposing it? In other words, if warrant for Q is a prerequisite for all beliefs based on perception, then Q cannot itself be warranted by means of perception. But it is hard to see how Q could be warranted any other way. Thus, there seems to be no way to acquire a warrant either for Q or for the more ordinary perceptual beliefs for which it functions as a presupposition.

The response to this problem on the part of the entitlement theorist is to account for how there can be warrant for authenticity conditions that is not acquired in the way that evidential justification is acquired. Wright defends an account according to which entitlement and justification are both species of epistemic warrant, but entitlement, unlike justification, doesn't require any specific evidential work to acquire. Entitlements are unearned in the sense that they are in place by default. If successful, this suggests a gap in the above sceptical argument. From the point that there is no way to *earn* a warrant for an authenticity condition, it need not follow that such propositions are unwarranted, so long as we can make sense of this idea that they can be warranted by default.

What is an entitlement a warrant to *do*? It doesn't simply follow from the possibility of a type of warrant that does not require evidence or other cognitive achievement that this would be a warrant to *believe* a proposition. A widely shared intuition is that belief is governed by some kind of evidential norm such that one ought only to believe propositions that are supported by one's evidence.[36] What is required, therefore, is a propositional attitude that is belief-like in some important sense, while lacking the strict evidential requirement of belief.[37]

[36] For an influential defence of this view see Feldman & Conee (1985).
[37] This particular strategy of appeal to a belief-like propositional attitude has been explored elsewhere. For example, see Pritchard's discussion of the non-belief reading of hinge propositions (2016, 90). Both Wright and

Wright thinks that there are other attitudes, other *modes of acceptance*, that capture the idea that one can, in a variety of senses accept a proposition in the absence of evidence (2004, 176). One can accept that P in a number of different ways that do not involve evidential support. For example, one can take P for granted for the sake of reasoning or act on the assumption that P in deliberation. In acting on the assumption that P I act as I would if I believed that P without actually believing that P. He settles eventually on the notion of trust on the grounds that, like belief but unlike *taking for granted* or *acting on the assumption*, trust it is incompatible with open-mindedness about the truth of the proposition in question. Anything less than rational trust in the truth of one's authenticity commitments will be "fully consistent" with agnosticism, which "seems impossible to square" with outright belief in the ordinary propositions which lie within the scope of those commitments (2004, 193). In rationally trusting that P one therefore adopts much the same kind of investment of confidence in the truth of a proposition as one would were one to believe it, taking P to be among those propositions that make up one's picture of the world. What distinguishes belief from trust, according to Wright, is the kind of epistemic support that each enjoys. Trust comes apart from belief "in cases where one is warranted in […] trusting that P for reasons that do not bear on the likely truth of P" - (2004, 177). Thus, though indistinguishable at the level of their *outward facing commitments*, i.e. their commitment to a proposition *as true*, belief and trust are distinguishable by their *inputs*, i.e. in the type of rational support they receive: while belief is warranted by evidence, trust is non-evidentially warranted. This raises the question of what it could mean for a proposition to be non-evidentially warranted, to which we will we now turn.

2.2 In Defence of Entitlement Theory

Wright explores a number of different ways to make sense of this idea of a non-evidential warrant to trust. Two of these ways get the most attention both from him and from various respondents in the literature. I will focus on those two for the purposes of this discussion.[38] Strategic entitlement is inspired by a Reichenbachian, decision-theoretic defence of the rationality of our practice of relying on induction. Entitlement of cognitive project, on the other hand, is inspired by Wittgenstein's remarks on scepticism. I take each in turn.

Pritchard give accounts of a belief-like attitude for propositions such as there is an external world, and so on. But whereas Wright's notion of trust is distinguished from belief via the evidential connection, Pritchard cashes out his account in terms of a lack of responsiveness to "rational considerations" (ibid). Unlike Pritchard, Wright does not think that the attitude we take towards these propositions are unresponsive to rational considerations, merely that the rational considerations in questions are not of the evidential kind.

[38] For Wright's discussion of entitlement of rational deliberation and entitlements of substance see 2004 (pp 197-203).

One strategy that Wright offers for the rationality of entitlement is inspired by Hans Reichenbach's pragmatic vindication of induction. Reichenbach was responding to Hume's inductive scepticism: we rely on induction, though we have no way of knowing that induction is, in fact, reliable. Reichenbach's master thought was that perhaps the practice of relying on induction can be justified if it can be shown that it is a dominant strategy—in the (decision-theoretic) sense that one strategy dominates the alternatives if, in pursuing it, one has nothing to lose and everything to gain. Consider four possible outcomes created by the following two variables: either we rely on induction or we don't, and either induction is reliable, or it is not.

	RELY ON INDUCTION	DON'T RELY ON INDUCTION
INDUCTION IS RELIABLE	MANY TRUE BELIEFS	FEW TRUE BELIEFS
INDUCTION IS NOT RELIABLE	FEW TRUE BELIEFS	FEW TRUE BELIEFS

Of the four possible outcomes, one is better and the other three are equally bad. This is what it means to say that we have everything to gain and nothing to lose by relying on induction. By relying, we open up the path to acquiring many true and useful beliefs. The only scenario that produces the desired outcome is the one in which we rely on induction and induction is reliable. In each of the three other cases, we have few true and useful beliefs; either because we rely on induction and form beliefs via inductive methods, but it turns out induction is unreliable, or we refrain from relying on induction and thus from no inductive beliefs at all. Induction is thus the dominant strategy in the sense that it can do no worse than the other strategies and can potentially do much better. Crucially, 'doing better' in this sense means doing better *from an epistemic point of view*. We already knew from Pascal-style reasoning that we can sometimes have good pragmatic reasons to believe in the absence of evidence. But what is important about Reichenbach's innovation, if it is successful, is that it now looks as though we can talk about reasons for belief that serve our epistemic goals while being themselves non-evidential.[39]

[39] It isn't clear that this decision-theoretic strategy is, ultimately, successful. See Daniel Elstein and Carrie Jenkins' 'The Truth Fairy and the Indirect Epistemic Consequentialist' for an argument as to how this sort of strategy does not result in properly *epistemic* reasons for belief (2020).

Wright's thought is that a similar argument could be given to account for the rationality of entitlement. Either one's perceptual faculties are generally reliable, or they are not. If they are, then trusting in their reliability opens up the path to acquiring perceptual evidence and justification. If they are not, then that path is closed no matter whether one trusts or not. So one has everything to gain from trusting in perception and nothing to lose. Thus, trusting is the dominant, and hence rational, strategy. Likewise, for other authenticity conditions too. One has everything to gain from rationally trusting in the existence of the external world, in the reliability of one's memory, and so on. In each case, the worst that can happen is no worse than if we do not trust, while the best that can happen is that we open ourselves up to the possibility of acquiring true beliefs.

Entitlement of cognitive project amounts to roughly the following idea: there is no coherent concept of rationality that is free from presupposition and hence the sceptic's demand that we provide justification for all propositions we accept is an incoherent ideal which we need not live up to. This model of entitlement is inspired by Wittgenstein's notes on scepticism, posthumously published as *On Certainty*. Wittgenstein writes:

> "Whenever we test anything, we are already presupposing something that is not tested… One cannot make experiments if there are not some things that one does not doubt… If I make an experiment, I do not doubt the existence of the apparatus before my eyes. … If I do a calculation I believe, without any doubts, that the figures on the paper aren't switching of their own accord, and I also trust my memory the whole time and trust it without reservation" – (*OC*, 163)

The point is that this practice of presupposing in the obtaining of certain epistemically-friendly conditions is, in a certain sense, unavoidable. One cannot help but take certain things for granted, not merely in a pragmatic sense but in a logical one. If this is true, then the sting is removed from any sceptical argument that demands we engage in a practice of inquiry free of presupposition: "If there is no such thing as a process of warrant acquisition", writes Wright, "for each of whose specific presuppositions warrant has already been earned, it should not be reckoned to be part of the concept of an acquired warrant that it somehow aspire to this—incoherent—ideal" (2004, 190). It merely seems as though it is some failure on our part to take on board certain implicit assumptions when engaging in inquiry, but this is no failure at all if doing so is a necessary feature of the concept of rationality. We cannot carry out our investigations without first accepting a set of assumptions about our cognitive abilities and the

reliability of our perceptual faculties that are themselves not the outputs of those abilities or faculties.

2.3. Classic Problems for Entitlement Theory: Alchemy and Leaching

As with the Moorean inference, many of the candidate entitled propositions are deductively entailed by the corresponding quotidian propositions. I believe that (P) here in front of me are two hands and this entails that (Q) there is a mind-independent external world. Now suppose the entitlement theorist is right such that I enjoy genuine evidential justification for my belief that here are two hands, partly in virtue of the corresponding perceptual evidence and partly in virtue of the background entitlement to trust that there is a mind-independent world. What is to stop me from recognising the entailment from P to Q and thereby, in the words of Martin Davies, transforming "the lead of rational trust into the gold of justified belief" (2004, 220)? As the wry metaphor suggests, such transformations seem too good to be true. Alchemy also seems at odds with the notion that entitlement theory represents a concessive response to scepticism. Part of the attraction to entitlement theory is that it can, at one and the same time, grant to the sceptic that cornerstones lay beyond the reach of evidential inquiry while mitigating the epistemological damage of this concession. But if alchemy is licenced by entitlement theory, then it appears we are in a position to acquire evidence for cornerstones and so there is no longer any concession to be made to the sceptic that cornerstones lay beyond the reach of inquiry.

A closely related problem to alchemy is the leaching problem. Here the worry is that the status of 'mere' entitlement for authenticity commitments threatens to spread upwards into any propositions that depend on these foundational hinges. As Wright himself puts it, "what prevents this 'merely taken for granted' character from leaching upwards from the foundations, as it were like rising damp, to contaminate the products of genuine cognitive investigation?" (2004, 207). If, at the foundation of our epistemic practices, lie assumptions that are merely taken on trust—however rational such trust can be understood to be—then nothing stronger than mere trust can be the output of the forms of inquiry that rely on such foundations. Another way that Wright has articulated this worry is in terms of a conditional: if there is risk in accepting Q, then there is risk in accepting any proposition for which Q functions as an authenticity commitment for the method of inquiring into whether P. And furthermore, if such risk precludes knowledge of Q, given that P inherits this risk, it surely then precludes knowledge of P also. The leaching problem is, therefore, at odds with the general goal of

entitlement theory. It seems as though the attempt to limit the sceptical import of the concession that we can claim no evidence for authenticity conditions cannot ultimately be contained in the way needed to sustain genuine evidential achievement for quotidian propositions.

Wright's response to this problem—or rather these twin problems—has evolved over time. Initially, he sought to respond by pointing out that both problems require certain formulations of closure that have already been discredited by the version of entitlement theory he defends. Recall that closure for evidential justification is jettisoned in favour of closure for his disjunctive notion of warrant. Whenever one has some warrant for P one ought, given closure for Wright's conception of warrant, have some warrant for anything entailed by P. But this is compatible with one having different kinds of warrant for P and its entailments. Wright motivates this point by reflection on the thought that one might have, say, a visual warrant for *those animals are zebras* that is in no way also a visual warrant for *those animals are not mules that have been disguised in some visually undetectable way* (2004, 229). And insofar as the leaching and alchemy problems seem to depend on some form of unrestricted evidential closure, they seem to miss the intended target.

More recently, Wright has changed his positions on these issues, in light of an argument from Aidan McGlynn which shows that placing the restrictions on closure that Wright advocates is not sufficient to block the alchemy and leaching problems. Up to this point, talk of closure has been specifically about closure across known or warranted deductive entailment. It is closure of evidential justification across entailment that Wright is in favour of abandoning. However, McGlynn shows that we do not need entailment closure in order to generate these problems. We can generate them with some extremely compelling alternative closure principles that, unlike entailment closure, so far no one has suggested abandoning. The revised argument appeals to two alternative closure principles. The first is what McGlynn labels:

> *ClosOR*
>
> If one has justification for P and one knows that P entails (P or Q) then one has justification for (P or Q).[40]

Whatever one thinks about the rejection of standard closure principles for entailment, rejecting *ClosOR* seems unthinkable. After all, to have evidence for a disjunct *just is* to have evidence—

[40] This principle appears in McGlynn (2012, 180).

the very same evidence—for the disjunction. But in combination with closure for a priori known equivalence, the alchemy and leaching problems return.

> *Equivalence Closure*
> If one knows a priori that P and Q are equivalent and one has justification for P, then one has justification for Q.[41]

Here is how the revised argument goes. Let P stand for *here is a hand*, Q stand for *there is an external world*, and assume that it is known that P *entails* Q. Next, assume that one has evidential justification for P. Now ClosOR tells us that one must have (the very same) justification for (PvQ) *here is a hand or there is an external world*. And since P entails Q, Q is an a priori equivalent of PvQ. Finally, since PvQ is equivalent to Q and one has justification for PvQ, equivalence closure tells us that one has justification for Q. Thus, the alchemical transformation of Q is complete.[42] Furthermore, it is not difficult to see how the argument can be run in reverse, starting out with the assumption that one *lacks* evidential justification for Q, going on to conclude that one therefore lacks justification for P, thereby reviving the leaching problem too.[43] This argument creates the same trouble for Wright's entitlement theory as before without appealing to the kind of entailment closure principles which Wright rejects. The two principles used here seem epistemologically watertight, as Wright himself acknowledges: "the two principles deployed […] may also seem ungainsayable" (2014, 232). While standard entailment closure principles seem intuitively very compelling, there is at least a history of other philosophers rejecting them in the face of paradox. There is no such precedent with respect to ClosOR and Equivalence Closure.[44] Wright does consider whether perhaps the correct response to the tension brought out by this kind of paradox is to reject *Equivalence Closure* after all (*ibid*.). However, though he briefly gestures in this direction, it is not the path he ultimately takes.

In light of the above problems for the rejection of evidential closure, Wright's response is to concede that there is both leaching and alchemy after all, but that these can be lived with.

---

[41] This principle also appears in McGlynn (*ibid*.).
[42] Alchemical because, of course, we are assuming in line with Wright's general approach that the original justification one has for P is conditional on a prior entitlement for Q.
[43] A structurally similar argument is given by John Hawthorne in his 2005 'The Case for Closure.'
[44] Granted, one might think this is simply because they have yet to be given the kind of extensive discussion in the literature that more standard equivalence closure principles have. Nonetheless, there is a clear sense in which equivalence closure and ClosOR seem even harder to give up than entailment closure.

On the face of it this seems very difficult to square, given that the upshot of leaching is that we cannot acquire evidential justification for either the ordinary everyday beliefs or the foundational hinge beliefs, while the upshot of alchemy is that we can have both. The way around this problem for Wright is first to limit the force of the alchemy problem by maintaining that, although we may indeed arrive at evidential justification for a proposition for which we started out with merely an entitlement, the rational credibility of that evidence is bounded by the prior rational credibility of the entitlement (2014, 233). One's epistemic situation "changes," Wright maintains, "but *without improvement*" (2014, 234).[45] The second key move is to countenance leaching but locate it at a higher order. That is, it is one thing to *possess* a warrant, and another thing to *lay a claim* to so possessing. And it is at the higher-order level of laying claim to a warrant where our warrants for so claiming are a matter of entitlement. By this notion of a claim to possessing warrant Wright has in mind "something to be assessed, and sustained or rejected, in a context of rational discussion and adduction of evidence, commonly recognized—very much as a claim to innocence, or guilt, may be discussed and assessed in the forum of a court of law" (2014, 220). And it is in the epistemological court room, as it were, that the moves we are permitted to make in assessing our epistemic situation are a matter of entitlement. Thus, leaching and alchemy co-exist but at the separate levels of, respectively, claims to, and possession of, warrant.

It is an interesting question whether leaching and alchemy can be kept apart at these different levels in the way Wright suggests. One worry is that perhaps the leaching problem will come back to bite once we recognise that our higher-order attitudes put certain constraints on our lower-order attitudes—as when, for example, I seem to be perceiving the world perfectly well but learn that I may have been given a dose of a drug which would cause me to misperceive. However, this is not the worry that I wish to pursue here. Rather, the rest of this chapter will be concerned with a novel problem for Wright's entitlement theory that arises once we consider the question of how confident we are warranted in being in a proposition on the basis of an entitlement.

2.4 A Question of Confidence

Plausibly, trust admits of degrees in a similar way to belief. Just as it seems quite natural to say that one believes P 'more', 'less', 'partially' or 'fully', it seems also quite natural to say that one trusts in P 'more', 'less', 'partially' or 'fully'. Moreover, since the claim is that trust and

[45] Emphasis in the original.

belief are both modes of acceptance which share in common their outputs, and that degrees of confidence are part of those outputs, it seems reasonable to say that to either believe that P or to trust that P is to invest some degree of confidence in P.

We can distinguish between two types of answer to the question of which is the correct degree of rational confidence to place in a proposition on the basis of an entitlement. An entitlement is either a warrant to be fully confident—i.e. certain—or it is a warrant to be less than fully confident. This distinction is exhaustive. Either an account licences certainty or it does not. Call any account that licences nothing short of certainty a *strong* account of ET. Alternatively, call any account of ET a *weak* account so long as it recommends a less-than-certain but nonetheless positive degree of confidence (where positive simply means more confident than not). In terms of credences, *strong* accounts licence credence 1, while *weak* accounts licence some degree below 1 and above .5.

The dilemma for trust-based accounts of entitlement is thus that the entitlement theorist must specify whether they have in mind a weak or a strong account of ET, but, as we shall see, both options face seemingly insurmountable problems. The problem for *strong* accounts will be relatively easier to appreciate, so we shall deal with this first before moving onto *weak* accounts.

2.5. A Problem for *Strong* Accounts

Recall the two main arguments offered in defence of ET, the *strategic* and *cognitive project* models. It is important to stress that these are not in any way *truth conducive*. Speaking of the former, Wright says that "What the Reichenbachian thought provides us with is one relatively clear paradigm of how such a commitment can be rational for reasons which do not impinge on the likelihood of the truth of the assumption in question" (2004, 182). Let us grant for the sake of argument that this strategy is successful such that it validates a positive attitude of trusting acceptance in the truth of certain propositions. Can it be that this strategy rationalises a strong notion of entitlement rather than a weak one? Given the admission that the kinds of arguments offered in defence of entitlement theory are in no way an indication of the likely truth of the propositions in question, it is hard to see how it could.

Here is an argument that demonstrates that it cannot. The entitlement theorist concedes that the strategic defence of entitlement theory, while demonstrating that it can be epistemically prudent to rationally accept certain propositions in the absence of evidence, this is in not taken to be any kind of indication of the likely truth of the propositions in question. For all that the

Reichenbachian defence has shown, we may yet end up in that square of the belief-matrix in which the beliefs that we form on the basis of an entitlement are, alas, false. I take it that the entitlement theorist concedes at least this much. Well then, another way to state this concession is that, even with an entitlement to trust that P in place, ~P remains an epistemic possibility. For example, consider the authenticity commitment *there is an external world* and consider an alternative hypothesis such as *idealism*. Suppose that we are entitled to trust that there is an external world on the basis of the kind of reasoning offered by the entitlement theorist. Does this mean that idealism is no longer an epistemic possibility? It is hard to see how it could given the concession that an entitlement to P is in no way an indication of the likely truth of P.

Next, notice that if ~P is an epistemic possibility, then the credence that one is warranted in having towards ~P cannot be 0. Furthermore, coherence demands that, insofar as one has any credence in a proposition and its negation, then one's combined credence in the two must sum to 1. Thus, if one's credence in ~P is higher than 0, then one's credence in P must be lower than 1. What this means is that the arguments offered in defence of the rationality of entitlement to trust cannot warrant credence 1. Even on the assumption that those arguments are successful in rationalising some positive degree of trusting acceptance of hinges, it looks clear from the fact that they leave open the possibility that those hinges are false that they cannot warrant credence 0 that they are false. Call this the *argument from epistemic possibility*.

Here is one way that a defender of the strong account of ET might try to resist this conclusion. The argument from epistemic possibility appeals to the thought that the denials of hinges remain epistemic possibilities in spite of our entitlement to trust in the truth of the corresponding hinges. One might object, however, that this is a kind of category mistake. Whether something is an epistemic possibility or not is an evidential matter. But entitlement theory is prefaced on the thought that hinges lay beyond the scope of evidential norms and constraints. By definition, hinges are such that evidence for them is not a rational ideal. Why should we not equally think that absence of evidence for the denials of hinges is not a rational ideal. And if so, then the fact that there is no evidence which can rule out such denials is beside the point; it is simply a misapplication of the concept of epistemic possibility—itself an evidential concept—to the category of hinges and their denials.

This objection fails on the grounds that there is no reason to think that, just because hinges cannot be supported by evidence, they cannot be defeated by evidence. Hinge epistemology maintains that warrant for hinges cannot be evidential because evidence can be acquired only in a context in which hinges are already warranted. But it does not follow from the fact that there cannot be evidence *for* hinges that there cannot be evidence *against* them.

On the contrary, one of the entitlement theorist's explicit conditions for an entitlement to P to be in place is that there are no reasons or evidence to think that P is false.[46] Entitlement theorists are not committed to the thought that there cannot be evidence against hinges, and indeed nor should they be. For example, consider the issue of perceptual reliability. Assume, in accordance with entitlement theory, that there is no way to acquire evidence that my perceptual faculties are functioning reliably, but that I am entitled to trust that they are. Then suppose that I discover that I have been given a drug that is going to cause my perceptual faculties to malfunction in some undetectable way. There seems to be no in principle reason for thinking that I could not acquire evidence for this and, furthermore, that once I have, a result would be that my entitlement to trust that my perceptual faculties are reliable gets defeated. Or consider our reliance on memory as a guide to past experience. For the entitlement theorist, this is again a matter of a non-evidential entitlement to trust. But suppose that I discover I am one of the test subjects in an advanced memory-implantation experiment. This discovery ought to give me sufficient reason to doubt something which I was previously entitled to take on trust, namely that my memories are an accurate representation of my past experience. What these examples bring out is that it does not follow from the fact that one cannot acquire evidence for a hinge one cannot acquire evidence against it. Thus, the objection to the argument from epistemic possibility under consideration fails. Even where a proposition cannot be supported by evidence and thus is a matter of a non-evidential entitlement to trust, it nonetheless still makes sense to think of its negation as an epistemic possibility and therefore our credence in its negation cannot be 0.

2.6. Problems for *Weak* Accounts (Anaemic)

A weak account is any account that licences degrees of confidence less than 1. However, I want to draw a further distinction between two kinds of weak account. One kind of weak account relates to degrees of confidence that are very high but just short of 1, while another kind relates to relatively lower (but still above .5) degrees of confidence. Call the former moderate and the latter anaemic.[47] Both accounts face serious problems. We shall begin with anaemic because,

[46] For example, this is condition (i) of Wright's account of entitlement to cognitive project (2004, 191).
[47] Martin Smith introduces the terminology of weak and moderate accounts of entitlement, although the distinction he makes is not the same one I make here (Smith, 2020). Whereas the distinctions I am here drawing relate to the strength of credence one may have in a proposition on the basis of an entitlement, the distinction Smith draws is, while not unrelated, specifically between accounts of entitlement that license either trust, belief or knowledge.

as we shall see, it is a more natural way to understand this notion of 'mere' trust as opposed to belief.

As noted earlier, a natural answer to the confidence question is that entitlements warrant mere partial degrees of confidence. Where I am entitled to trust that P, I am entitled to be more confident of P than I would if I were merely agnostic, but not so confident as to count as believing P, and certainly nowhere near being certain of P. This answer fits neatly with speaking about this species of warrant as "mere" entitlement (Wright, 2004, 208). It would be great if we could have evidential justification to *believe* that there is an external world and so on, but in the absence of such evidence, perhaps an entitlement to merely trust—in this weak sense—in the existence of the external world will suffice. This way of thinking suggests that we should look to *anaemic* accounts of entitlement for our anti-sceptical strategy.

I will now provide two problems for *anaemic* accounts of entitlement. It will be clear that both problems are very much in the same spirit as one another but pertain to two different types of hinge proposition. The first problem concerns some probabilistic constraints across deductively valid arguments, while the second concerns an epistemological principle that constrains degrees of confidence that rational subjects are warranted in placing in the output of epistemic methods or faculties.

2.6.1. The Probabilistic Argument.

The first problem for anaemic accounts concerns the set of hinges for which it is the case that there is an entailment from the ordinary proposition P to the hinge proposition Q.[48] Given such an entailment relation, considerations from probability theory will force a restriction on the rational confidence an agent may place in the ordinary proposition given a restricted degree of confidence in the hinge.

Consider the following consequence of the probability axioms:

$$\text{If } P \rightarrow Q \text{ then } Pr(P) \leq Pr(Q)$$

[48] See section 2.8. for a case in which there is no entailment from the ordinary proposition to the hinge proposition q.

Where P entails Q, the probability (*Pr*) of P cannot be higher than that of Q. This is a principle about objective probability. However, notice that, on assumption that a rational agent's credences will satisfy the laws of probability,[49] we can derive the following related principle:

> (RC) where P implies Q, the credence that it is rational to assign to P must be a least as high as the credence that it is rational to assign to P

This principle is eminently plausible. A simple example may help to show why. The proposition this is a red ball deductively entails that this is a ball. RC says that it would be irrational to be more certain that this is a red ball than that this is a ball. And that seems clearly right. By carrying out a deductive inference, one is guaranteed not to introduce any new error possibilities. This is why, as Richard Pettigrew frames the same point, a rational agent's credences should not 'drop' across an entailment.[50] The principle appears intuitively very compelling, at least with respect to the notion of an ideal rational agent. Real-world subjects will sometimes rationally violate RC for the reason that we are not logically omniscient and fail to recognise the deductive consequences of some of our beliefs. In what follows, we will bracket this concern by assuming that in each case, the agent is certain of the entailment.

Taking this principle about credences as our starting point, we are now able to construct a puzzle. First assume that P implies Q and that Q is a hinge proposition for P.

The degree of confidence that it is rational to assign to Q is at least as high as the degree of confidence that it is rational to assign to P	(RC)
I am warranted in having high confidence in P	(*optimistic ass.*)
I am warranted in having (at most) low confidence in Q	(anaemic ET)
I am warranted in having high confidence in Q	(1,2)
⊥	(3,4)

[49] This assumption is known as probabilism. For a discussion of probabilism and related principles, see Pettigrew (2016).

[50] Pettigrew labels a version of this principle "No Drop", which, loosely stated, says that "rationality requires that an agent's credences not drop over a logical entailment" (2016, 2). More formally, No Drop says that "If an agent has opinion set {A, B} and A entails B, then rationality requires that $c(A) \leq c(B)$" (ibid). While some readers may object to the idea of capturing our epistemic practices using the axioms of probability, I presume that such readers will still appreciate there is a sense in which it is irrational to be more confident of the antecedent of a conditional than of its consequent (assuming that one grasps the entailment).

The assumption at step 2 is one that all (non-sceptical) epistemologists ought to grant. Call this the *optimistic assumption*. However, when we combine the *optimistic assumption* with the RC principle, we get the thought that I must be warranted in having a high degree of confidence in Q. This is in direct contradiction with premise 3, which is the statement of anaemic ET. Thus, we have our contradiction. The anaemic reading of ET when applied to the kinds of cases where there is an entailment from the quotidian proposition to the entitled hinge proposition is incoherent.

One response that we might anticipate the entitlement theorist making to this argument is very similar to the kind of position Wright takes in response to McGlynn's revived evidential closure argument which seemed in turn to revive epistemic alchemy. Wright's considered view seems to now be that there is alchemy after all, but that this can in some way be lived with. If we assume for the moment that alchemy is permissible, then it looks like we have a way to push back on premise 3 of the above argument. What our imagined respondent might say is that, indeed, on the basis of an *entitlement* the most one can have is a low degree of confidence in Q, but that given evidential closure and the alchemical transformation it licences, one can then end up with actual evidence for Q and on the basis of this evidence one can have a much higher degree of confidence in Q. If this is correct, then premise 3 fails.

What our imagined respondent gets wrong here is that the admission of alchemy does not come without caveats. Although Wright is willing to grant the possibility of acquiring evidential justification for hinges via transmission, he reminds us that, nonetheless, "the significance rationally assigned to the evidence concerned is bounded by one's anterior confidence in the conclusion, so that the argumentative routine in question is not available to enhance that prior confidence" (2014, 234). For example, contra his earlier account of entitlement theory and the failure of transmission of warrant that goes along with it, one may be able to reason through an inference and acquire evidence to believe that, say, *those animals are not cleverly disguised mules*. However, "the strength of the visual evidence against the disguised mule possibility is bounded by the prior credibility of the proposition that it does not obtain" (2014, 235). It's a kind of alchemy then, but not one that serves up a credence-boost. Even granting the alchemical transformation of Q, premise 3 seems to hold up after all.

2.6.2. The Epistemological Argument

This argument is related to the previous but is logically distinct. The trouble highlighted in the previous case will arise for all hinges that are deductively entailed by the propositions for which

they function as hinges. This is true of many hinges. For example, the proposition *I have hands* deductively entails *I am not a handless brain in a vat*. However, it is not always the case that hinges are deductively entailed by the propositions for which they function as hinges. Consider, for example, the following pair:

(P) I have hands
(Q) my perceptual faculties are functioning reliably

There is no deductive entailment here from P to Q. It is perfectly compatible with the hypothesis that I am the victim of a sustained hallucination that I have hands. Nonetheless, on the assumption that my warrant for believing that I have hands is perceptual, then Q is still a hinge for P and therefore among those propositions to which I need an entitlement to trust. The argument from the previous section will not apply in this case, given that the RC principle applies only to those cases where there is a deductive entailment between the ordinary and hinge propositions. There is, however, a fairly plausible epistemological principle that will give us more or less the same result. Suppose that one wants to carry out an investigation. There is a method and an output or answer that the method will produce. Very plausibly, the confidence that one places in the reliability of a method places an upper bound on the confidence that one may rationally place in the output of that method. Call this the *Upper-Bound Principle*:

(UBP) a rational subject S cannot be more confident in the output of an epistemic faculty or method M than S is in the reliability of M

UBP is eminently plausible. If one wanted to test the temperature of a liquid and the only method one had of doing that was to use a thermometer, it would be irrational to be less confident in the reliability of the thermometer as an instrument for detecting temperatures of liquids than in the reading one got from it. Generally speaking, UBP applies to empirical beliefs formed via perception too. For most of the propositions about which we form beliefs via perception, perception is the only method of inquiry that we have. Thus, for beliefs in those propositions UBP is applicable and we cannot rationally be more confident in them than we can in the generally reliability of perception itself.[51] Note, however, that UBP will apply only

[51] For some beliefs formed via perception, it might be the case that we can have independent, non-perceptual reasons or evidence for them and therefore this argument will not apply. Consider the well-known anecdote from the history of science which has Galileo reasoning from the armchair that objects fall to the ground at the

to cases in which subjects have some degree of confidence in the reliability of a method or faculty—the issue is one of incompatible combinations of degrees of confidence, so insofar as one has no degree of confidence in the reliability of a method or faculty, there can be no incompatibility.

Wright himself seems to find something in the ballpark of this principle, not merely plausible, but "barely more than a platitude" (1991, 99). The principle that he introduces is the following.

> *Proper Execution Principle (PEP)*
> If the acquisition of warrant to believe a proposition depends on the proper execution of some procedure, then executing the procedure cannot give you any stronger a warrant to believe the proposition in question than you have independently for believing that you have executed the procedure proper.

Though PEP is not equivalent to UBP, it is nonetheless very much in the same spirit and will be just as effective in generating the puzzle below. A couple of differences are, firstly, that UBP is a principle about rational confidences while PEP is a principle about warrant. However, Wright mentions in a footnote that a corresponding principle about 'reason to believe' is "just as plausible" as PEP. Secondly, UBP makes no mention of having *independent* warrant or reason to believe in the reliability—or proper execution—of one's belief-forming method. Though Wright takes PEP to be a 'platitude', this requirement that one have *independent* warrant or reasons to believe in the reliability of a method or procedure is very clearly a *conservative* requirement, which there is at least conceptual space for a liberal about belief-forming methods to disagree with.[52] At least this has to be so if we think that a claim that is contested by sincere interlocutors cannot be a platitude. Presumably, the real platitude is that the reasons for belief that one gets by virtue of executing the procedure are capped by the degree to which one is confident that one has executed the procedure properly. If so, the principle is effectively equivalent to UBP.

---

same velocity irrespective of differences in mass. Assuming his reasoning was sound and that one *could* also have perceptual evidence for the same conclusion, it seems that one could have a belief based on perception of which one was more confident than one was in the reliability of perception. However, this would not, strictly speaking, be a counterexample to UBP since the belief would only be *partly* based on perception.

[52] For an account of how a liberal can reject the independence claim, see the proposal I offer for how the reliabilist ought to deal with the *Coins* case in chapter 3 section 3.5.

With UBP on the table, we are now able to appreciate another puzzle generated by UBP together with ET and the *optimistic assumption*. Let Q be *my perceptual faculties are functioning reliably* and P be a proposition that one is warranted in believing solely on the basis of perception such as *here is a hand*.

One cannot be more confident in P than in Q	(UBP)
I am warranted in having high confidence in P	(*optimistic ass.*)
I am warranted in having (at most) low confidence in Q	(anaemic ET)
I am warranted in having high confidence in Q	(1,2)
⊥	(3,4)

The *optimistic assumption* and UBP jointly entail that I am warranted in having a high degree of confidence in Q. But, as before, this is in direct contradiction to the anaemic reading of ET. Insofar as we find UBP and the *optimistic assumption* plausible, the puzzle can only be solved by rejecting the anaemic reading.

The lesson from this puzzle is that an entitlement needs to be a warrant for being at least as confident in the presuppositions of a particular project or method of inquiry as in the outputs of that method. If we want to allow that subjects can be warranted in placing high degrees of confidence in propositions concerning the presence of objects in one's immediate visual field, then the entitlement theorist needs to account for how we can be equally confident of the reliability of one's perceptual faculties delivering the visual information about such objects.

What these two problems show is that an entitlement theory that is weak in the *anaemic* sense outlined above will not be sustainable. More accurately, it will not be sustainable as long as it is intended to fit with high degrees of confidence in ordinary everyday propositions. One way for the entitlement theorist to respond to these worries might be to argue that the degree of confidence that is licensed by entitlement theory is in fact very high. Not so high as to count as certainty, but near enough. If we are entitled to very high—albeit non-maximal—degrees of confidence in hinge propositions, then we can sometimes also be justified to hold equally high degrees of confidence in ordinary propositions. And maybe this is enough to be getting on with. It would be nice if we could be certain some of the time but if the upshot of philosophical investigation is that certainty is always beyond our grasp, then this need not be such a damaging

realization if we can still be very highly confident at least some of the time. However, as we shall now see, even this 'moderate' version of ET will be unstable also.

2.7. Problems for *Weak* Accounts (Moderate)

The moderate account says that an entitlement to trust is a warrant to be very highly confident of hinge propositions, though it stops short of warranting certainty. Dialectically speaking, this account looks like a promising move for the entitlement theorist, in that it sidesteps the problems identified for both the anaemic and strong accounts considered above. So long as we can be content with capping our degrees of confidence in propositions at a high but less-than-full degree, we will not be forced into condoning the kind of irrational doxastic attitudes that caused problems for anaemic accounts. Granted, this means that certainty with respect to both hinges and ordinary propositions is forever out of reach. But perhaps this is not too huge a price to pay given the threat of scepticism.[53] Likewise, this view will not suffer from the kinds of challenges faced by strong accounts of entitlement. The inherent conflict between outright certainty in hinge propositions and the admission that it is epistemically possible that one is in error with respect to such propositions, goes away once we give up on outright certainty. There is clearly nothing incoherent in affirming that one is almost certain that P is true and yet that ~P is an epistemic possibility. This moderately more robust theory of entitlement therefore looks more promising than both the anaemic and strong accounts.

It is clear, however, that this position is in fact utterly unstable. To appreciate why, first notice that for any one proposition that enjoys evidential support, there will not be merely one hinge proposition in place but rather arbitrarily many. And the crucial point is that small degrees of doubt about individual hinges will lead, when those hinges are conjoined together, to high degrees of doubt about their conjunction. Consider again the 'ordinary' proposition *there is a laptop on the table in front of me*. In discussing this proposition earlier on we took the example of *my perceptual faculties are functioning reliably* as a respective hinge assumption. But it is easy to see how to inflate the relevant set of hinges. Not only does one need entitlement for this hinge assumption and for the other usual suspects—*I am not a brain in a vat; I am not dreaming; I am not in a Matrix-style scenario; there is an external world to which I have perceptual access*—but one would need warrant to discount each possible

[53] The price will be much higher for one who is already committed to some forms of infallibilism which says that if one knows that p then one is certain that p. On the assumption that we are committed to infallibilism about knowledge, all versions of ET will have the consequence that knowledge—of hinges an of ordinary propositions—is impossible.

sceptical scenario, and the set of all possible sceptical scenarios is going to be arbitrarily large, given that there seems to be no principled limit to the number of distinct sceptical scenarios that one could be in at any one time. This argument assumes that distinct sceptical scenarios represent distinct possible worlds that one could be in at any one time which would entail that one has the same subjective experience that one currently has. So, in one sceptical scenario I have been envatted and tricked into believing that I am having the experience that I currently take myself to be having for the purposes of a late-night television show, while another scenario differs from this only in that I have been envatted for the purposes of a university laboratory experiment. If this characterisation is correct, then it is easy to appreciate how a set of possible sceptical scenarios can be made arbitrarily large.

One might object that these scenarios are not really distinct qua *sceptical* scenario, insisting that the differences between the two are superficial and do not really pertain to what is pertinent to them as hypotheses about the causal origin of my experience. This objection would insist that two sceptical scenarios that are purportedly distinguishable by arbitrary non-essential features are not really distinct scenarios. While I have some sympathy for this line, I find it very difficult to see on what principled grounds one could distinguish between superficial and pertinent details of scenarios. In the absence of such a principled distinction, I will take each possible world compatible with a subject's subjective experience yet incompatible with what they believe systematically to be one sceptical hypothesis.[54] Thus, there is no in principle limit to the number of distinct sceptical scenarios compatible with a subject's subjective experience.

Now, with respect to a given ordinary proposition and its evidential basis, let C^1 be a hinge proposition, and let C^2 be an alternative sceptical hypothesis that is incompatible with C^1, likewise, for C^3, C^4, and so on. The set of all possible competing hypotheses is then ($\forall$C): $C^1+C^2+C^3 \ldots C^n$. Suppose that we can indeed make 'n' arbitrarily high. Next, consider the question of how our credences should be spread across the set $\forall$C. According to the moderate account of ET, we are warranted in being, at most, highly confident in those propositions we are entitled to trust such as C^1. And suppose furthermore that, as the *moderate* account holds, the degree of confidence that we are entitled to have towards any given hinge (C^1) is less than 1, which in turn means that we are rationally committed some positive credence to C^2, and again some positive credence to C^3, and so on. Now, the problem that we face is that however

[54] The general set may be domain-specific, as for example sceptical hypothesis about the existence of the past will be in conflict with one's beliefs about the past but not in generally in conflict with one's beliefs about the current state of the world.

low we make the rational warranted degree of confidence in each of the alternatives to C^1, the fact that there is no in principle limit to the number of these alternatives, means there is going to be very little credence left over to assign to C^1. To be clear, the credence that must be assigned to C^1 is no lower than the credence assigned to any of the other competing hypotheses, but this is no consolation given that the credence that must be assigned to any proposition in the set is now arbitrarily close to 0. The moral of this particular story is that on the assumption that the warranted credence in a hinge is less than 1, by reflecting on the fact that there is no in principle limit to the number of competing hypotheses, we find that the actual credence that is warranted in the hinge is close to 0.

2.8. Rejecting the Credence Game

So far, I have shown that once we allow for rational trust to be cashed out in terms of credences, we run into trouble when trying to answer the question of precisely what credence we are rationally permitted to take on the basis of an entitlement. A possible way to avoid the foregoing problems is therefore to reject this whole idea that trust admits of degrees. In section 2.1. I argued that trust admits of degrees on the basis that belief admits of degrees and that belief and trust are identical in all but 'input', i.e. the type of rational support they receive. But this argument will fail if this difference in input could conceivably give rise to a difference in whether these modes of acceptance are gradable. Here is how such an objection might run. The rationality of your attitude to hinges is not a function of your evidence. But the rationality of your credal attitudes *is* a function of your evidence. Ergo, your attitudes towards hinges should not be understood as credal attitudes.[55] By way of example, consider a set of hinge propositions such as (EW) *there is an external world*, (IN) *induction is reliable*, and (PR) *perception is generally reliable*. Entitlement theory predicts that we trust each of these equally. Moreover, if we consider their negations, ~EW, ~IN, ~PR, it seems like we deny these equally also. So, the claim made earlier on in section 2.1. that we ought to think that rational trust in hinges is gradable is not well supported by reflection on the ways in which we do trust them. Contrast this to ordinary, empirical propositions such as (CO) *there are cars outside*, (RO) *it is raining outside*, or (SO) *the corner store is currently open*. These look like the kinds of propositions that are a straightforwardly a matter of degrees of confidence. We can be certain that CO is true, fairly confident that RO is true, and not very confident that SO is true. In each case, where

[55] I would like to thank an anonymous journal referee for raising this objection.

our attitude is a matter of credence, what credence we should adopt is a function of our evidence, and thus credal states are inextricably linked to evidence.

By way of response to this kind of argument, the first thing to note is that, strictly speaking, all that we can really rule out about propositions such as EW, IN and PR is any variation in our rational confidences towards them. But we cannot infer from the intuition that we trust in each equally, that such trust is no kind of degree. Granted, with respect to credences of the everyday kind, it is easy to see how credences are fixed by evidential probability, and thus in the absence of evidence it is very hard to see how to rationally assign credences. This is a difficult problem, but it is not clear that it can be solved by simply stipulating a form of trust that does not involve any degree of confidence whatsoever. After all, it is compatible with the admission that credences are ordinarily a function of evidence that, in the extraordinary case, there is a given credence that we are rationally committed to take towards hinge propositions that is not a function of our evidence but that is axiomatic of rationality in the way suggested by the entitlement theorist, the above dilemma notwithstanding. Although this involves revising our orthodox conception of how credences function, it is arguably no more revisionary than the general thesis of entitlement theory which says that we can sometimes be warranted in accepting as true certain propositions in the complete absence of evidence.

A second thing to say in response to the above argument is to grant the point that (the relevant notion of) trust is *not* gradable and see where it takes us. Suppose that where we are rationally committed to trusting that P on the basis of an entitlement, this is not to be understood in terms of investing any degree of confidence in P. The problem this then raises is analogous to the problems raised for weak accounts of entitlement earlier on in section 2.7, namely what is the correct rational response to the recognition that a hinge proposition Q follows deductively from an empirical proposition in which one places a high degree of confidence? Rationality demands that where one is *a priori* certain that *P entails Q*, one must be at least as highly confident that Q is true as one is that P is true. For all that the entitlement theorist has offered by way of an appeal to trust as a mode of acceptance that is not a function of evidence, we are yet to hear anything which would suggest inferences involving entailments to hinge propositions are exempt from these kinds of constraints on rationality.

2.11. Concluding Remarks

This chapter began with a question for the entitlement theorist concerning the warranted degree of confidence licenced by entitlement theory. We noted that either an account licences certainty

or it does not. We have seen that accounts that seek to construe entitlement as warranting certainty are not at all supported by the two main arguments offered in defence of entitlement theory. Furthermore, we have seen that accounts that licence less than certainty lead, in various ways, to trouble. In order to demonstrate this, we have drawn a further distinction between two kinds of less-than-certainty accounts: anaemic and moderate. Anaemic accounts face the problem that, when combined with high degrees of confidence in ordinary propositions, they lead to irrational doxastic attitudes. Yet moderate accounts face the problem that even the smallest degrees of doubt about many hinge propositions multiply together leading to extremely low degrees of confidence in the complete set of hinges. No account of entitlement, be it weak or strong, is both coherent and well supported by the kinds of arguments entitlement theory appeals to.

In the absence of a solution to the dilemma proposed here, the entitlement theoretic response to the question of how there can be warrant for authenticity conditions seems to be in trouble. Entitlement theory is, broadly speaking, an internalist account of warrant: entitlements are in place irrespective of facts about one's environment, the functioning of one's faculties, or modal factors. In light of the problems for entitlement theory, it is worth considering whether externalism could offer an alternative. One obstacle to such an account is that it is often claimed that externalism is somehow incapable of offering a satisfying picture of transmission failure. The goal of the following two chapters is to challenge this claim.

3.

3. The Good, the Bad, and the Controversial: Externalism about Transmission Failure

Two common claims in the literature on transmission failure are that

(i) there is a general phenomenon of transmission failure which externalist theories of justification have difficulty in capturing

(ii) the answer to the question of what is wrong with the Moorean inference is that it exhibits transmission failure

In this chapter it is shown that the externalist is able to reject both claims. Firstly, reflection on a spectrum of candidate transmissive and non-transmissive inferences shows that, while the literature often focuses on more controversial cases of transmission failure, there are in addition a range of more clear-cut cases. Basic externalist theories (safety, reliability) seem to have no problem in capturing intuitions about the clear-cut cases of non-transmissive inferences, which suffices to reject (i). It is then discussed whether the externalist ought to reject (ii). Two pre-existing attempts to reject (ii) from within an externalist framework are considered, but it is shown that each account faces a range of problems. It is then argued that it is open to the externalist to deny that there is transmission failure in the Moorean inference and to appeal instead to the notion of what I call negative epistemic dependence to explain the source of the intuition of impropriety.

3.1. The Externalist Challenge

As noted in the first chapter, certain inferences are trivially non-transmissive such as the obviously question-begging inference *P, therefore P*.[56] One cannot acquire warrant for a proposition for which one does not already have a warrant by inferring it from itself. A more substantial example is the following. By counting the number of coins on the table as seven, I can acquire warrant for the proposition *there are seven coins on the table.* Two propositions are then warranted for me: *I just counted exactly seven coins on the table* and *there are seven coins on the table*. From here, there is a straightforward deductive inference to the proposition *I did not miscount the coins*. However, many who contemplate this inference find it to be

[56] Technically this assumes an additional premise to the effect of *P entails P*.

epistemically circular in some way. Although I do seem to be able to acquire warrant for believing that there is X number of coins on the table on the basis of counting, this warrant in some sense already presupposes that I have not miscounted, so that I cannot then warrantedly deduce that I have miscounted, on pain of a kind of epistemic circularity[57].

One notable feature of the literature on transmission failure is that the most dominant voices have tended to favour some form of internalism about warrant.[58] Taking Crispin Wright and Jim Pryor as, respectively, the foremost proponents of conservatism and liberalism, with Annalisa Coliva defending a possible middle ground 'moderate' position, the landscape of this debate is predominantly one in which warrant or justification are understood along internalist lines.[59] One potential point of confusion in labelling the contrasting views in this debate as 'liberal' or 'conservative' is that all parties should agree that we ought to treat some inferences liberally and likewise some inferences conservatively. Recall that to treat an inference from P to Q liberally is to hold that the acquisition of warrant for P need involve no prior warrant for Q. In contrast, to treat an inference conservatively is to hold that prior warrant for Q is part of what accounts for there being warrant for P in the first place. An inference that has the structure that conservatives say it does will therefore fail to transmit warrant from premise to conclusion: one cannot acquire warrant for Q via inference from P if the warrant for Q is epistemically prior to the warrant for Q in the way that the conservative says it is; nor will Q be warranted *in virtue of* the inference from P. Much of the discussion between liberals and conservatives is around whether to treat the Moorean inference conservatively or liberally. The kind of view that I will be advancing below is liberal in this sense. However, unlike the dominant accounts of liberalism in this literature it will appeal to externalist conceptions of justification in order to get this result.

While it is not difficult to see how an externalist account of justification will predict that many inferences are transmissive of justification, it is often suggested that it is hard to see how externalism can countenance transmission failure at all. Insofar as we have the intuition—independently of our commitments to internalist or externalist theories of justification—that there is an interesting phenomenon of transmission failure, it looks bad for the externalist if they are unable to account for it. And yet this is exactly what many in the debate take to be the

[57] This intuition is widely shared, though not universally. Bergman and Kornblith defend reliabilist accounts on which such 'bootstrapping' reasoning is permissible (Bergmann, 2006, Kornblith 2002). However, in section 3.2. I switch focus to a different case of transmission failure which does not depend on bootstrapping. Thus, even those like Bergman and Kornblith who think that the *coins* case does not exhibit transmission failure will be able to acknowledge transmission failure elsewhere.

[58] Two recent exceptions to this are considered at length below.

[59] See (Wright, 2000, 2002, 2004, 2012, 2014), (Pryor, 2000, 2004, 2012), (Coliva, 2012, 2015)

case. Wright himself is pessimistic about the hopes for accounting for the phenomenon on an externalist framework. He writes: "My own suspicion […] is that providing a convincing, recognizably externalist diagnostic of what is going wrong in at least a decent cross-section of the usual putative examples of non-transmissive arguments, is none too easy a thing to do" (2012, 455-56). The worry that Wright and others have seems to be that, since deductive inference is a paradigm reliable process, it seems difficult to see how there can be transmission failure on a reliabilist account of knowledge. Likewise for a modal account such as a safety condition. If modal robustness, in the sense of a belief that is true in all or most nearby worlds, is all that matters for knowledge, then how can such an account make sense of transmission failure given that (competent) deduction, by definition, cannot introduce any new error possibilities. "Externalism", Wright concludes, "may be forced to deny the phenomenon" of transmission failure altogether (*ibid*., 458). So much the worse for the externalist.

Similarly, Jose Zalabardo acknowledges that "it is often assumed that one can only [account for transmission failure] by imposing internalist constraints" [on warrant] (2012, 321). Zalabardo takes this to suggest a dilemma: either we accept such constraints, or we allow for unrestricted transmission. Unrestricted transmission looks like a very bad result, given the existence of clear-cut cases of transmission failure. The challenge for the externalist, therefore, is to reject the first horn of Zalabardo's proposed dilemma by arguing that one can accommodate the phenomenon of transmission failure without accepting these internalist constraints on warrant.

The goal of this chapter is to demonstrate that, contrary to the above remarks, there are ways that the externalist can capture the phenomenon of transmission failure. In one sense, this goal is not quite as ambitious as it sounds given that there are already a handful of externalist accounts on the market. However, I will argue that there are good reasons to think that these accounts face a number of problems, including failing to meet the minimal desideratum that I will lay out in the following section. Before getting to these, I will first discuss a variety of cases of alleged transmission and transmission failure in order to clarify which cases are crucial to the viability of any particular theory and which are not. By 'crucial' I mean it being a requirement of any plausible theory of transmission failure that it count them as non-transmissive. The thought is going to be that some cases are clear-cut cases of transmission failure and others clear-cut cases of successful transmission, while still other cases are more controversial, and that the standard by which we should judge whether a theory is able to capture the phenomenon ought to be, minimally, that it capture intuitions in the clear-cut cases. If our theory of transmission failure can get that much right, then we have a viable theory. A

second question will then arise as to what that theory says about the more controversial cases, however we should not presume that the theory stands or falls by the predictions it makes in those cases.

3.2. Some Cases (the Good, the Bad and the Controversial)

Call any case that is a clear and obvious case of perfectly acceptable warrant transmission a *good case*. By way of example, here is one of Wright's paradigm good cases.[60] Suppose that I learn that my friend has just eaten a large helping of risotto that contained a poisonous toadstool. There is no chance of medical intervention since we are several days travel from the nearest village with no means of communication and without the specialised medical knowledge or medicine to revert disaster once the poison has been ingested. I then reason as follows, given that my friend has ingested a lethal dose of toxins, he will shortly die.

Toadstool

(E) Jones has just consumed a large risotto of Boletus Satana

(P) Jones has absorbed a lethal quantity of the toxins that the toadstool contains

(Q) Jones will shortly die

The acquisition of warrant for P based on the kind of evidence in E need involve no prior warrant for Q. This argument is therefore well suited to transmit warrant from P to Q. One who acquires warrant for P via the kind of evidence given by E, and recognises the inference to Q, can thereby *acquire* warrant for Q.[61] The eventual warrant that one has for Q is then had *in virtue of* inference from warranted premises. This inference ought therefore to satisfy the kinds of transmission principles we discussed earlier on in chapter one.

Likewise, call any case a *bad* case if it is very clearly a case of transmission failure. Perhaps the most clear-cut case is the obviously question-begging inference, P, therefore P. It should be fairly obvious that blatant question-begging inferences exhibit transmission failure, almost as a matter of triviality. Wright, however, offers a different example as a clear-cut case of transmission failure, and we will focus on this one as our paradigm *bad case*.[62] Suppose a background context in which Sarah and Jessica are identical twins and suppose furthermore that I have great difficulty in telling them apart based on looks. I see someone ahead who looks

[60] The example originates in Wright (2002).
[61] For a similar case, see '*Betrothal'* in Wright (2002, 333)
[62] Again, the example originates in Wright (2002).

just like Sarah, form the corresponding belief that *this person is Sarah*, and finally deduce that therefore *this person is not Jessica*.

Twins
(E) this person looks just like Sarah
(P) this person is Sarah
(Q) this person is not Jessica

Do I have a warrant for P in this case? Clearly the answer is *no* if all I base my belief on is the mere fact that the person ahead looks like Sarah. In order for me to have a warrant to believe P based on appearances, I would need to already have warrant to rule out the possibility that the person is Jessica. Suppose then that I have such a warrant—perhaps a friend has earlier on informed me that Jessica is out of town on a business trip. Now it looks like the visual appearance, together with the background information, does provide me with a warrant to believe P. However, it should be clear that this warrant now will not transmit to Q. It is only because I already had warrant to believe Q that I could acquire warrant for P. Thus, inferring Q from P would not be a way to *acquire* a warrant for Q.

Toadstool and *Twins* represent the most clear-cut cases. It is relatively uncontroversial that warrant transmits in the good cases but fails to do so in the bad cases. If we conceive of all possible cases that fit this tripartite structure (E-P-Q) as on a spectrum from the most uncontroversial cases of successful transmission at one end to the most uncontroversial cases of transmission failure at the other end, then there are going to be many cases in between for which the status of transmissiveness is controversial. There are going to be varying degrees of controversy for different types of cases, with the most controversial ones at the centre of the spectrum. Call any case *controversial* if it belongs in or close to the centre of the spectrum.

Our paradigm controversial case is the familiar Moorean inference. Once again, it fits into the familiar E-P-Q structure:

Moore
(E) it appears as though I have two hands
(P) I have two hands
(Q) the external world exists

Is *Moore* more similar to *Toadstool* or to *Twins*? As discussed in chapter I., the answer to this question is disputed. Liberals like Pryor think that *Moore* represents a perfectly good route to acquiring knowledge of the external world,[63] while conservatives like Wright deny this.[64] All parties to the dispute agree that, to use Wright's term, Q is an authenticity condition for believing P on the basis of E: if one had cause to doubt Q, this would serve to defeat the justification that E would otherwise provide for P. But conservatives and liberals will disagree as to whether more than merely an absence of grounds for doubt is needed. If one takes a conservative view towards *Moore*, then warrant will fail to transmit from P to Q—one can't *acquire* a warrant for Q via inference from P if warrant for Q is among the preconditions that need to be in place in order to acquire warrant for P. On the other hand, on a liberal framework, one can acquire a warrant for a perceptually basic proposition such as the premise of *Moore* independently of having acquired warrant for its conclusion and thus it will be possible to acquire warrant for that conclusion via recognition of the entailment. While it is clear that one ought to be conservative regarding *Twins* and liberal regarding *Toadstool*, in conceiving of all possible cases exhibiting the familiar tripartite structure as on a spectrum, the difficult questions are going to be where to place other cases, such as Moore, about which there is more disagreement.

Consider the *Coins* case discussed briefly in section 3.1:

Coins
(E) I counted seven coins on the table
(P) there are seven coins on the table
(Q) I did not miscount

While a majority of epistemologists view *Coins* as on a par with *Twins*, there are some who are willing to treat it liberally in order to avoid what they see as an inevitable scepticism that would be the result of an overly conservative epistemology.[65] *Coins* perhaps fits somewhere in between *Moore* and *Twins* on our spectrum of cases.

Another important case in the literature on transmission failure is Dretske's zebra case.[66] Wright's conservative view, as discussed in chapter one, is that this case indeed exhibits

[63] For example, Pryor (2000, 2004, 2012).
[64] For example, Wright (2002, 2003, 2014).
[65] As mentioned above, see Kornblith (2002) and also Bergmann (2012), both of whom are happy to endorse the kind of reasoning exemplified by *Coins*.
[66] Dretske (1970, 2005)

transmission failure. In acquiring warrant to believe that *those animals are zebras*, one is antecedently committed to presupposing that *those animals are not cleverly disguised mules*.[67] It appears as though Pryor leans towards a conservative treatment of *Zebra* as well. He thinks that while the premise of *Moore* is immediately warranted, warrant for the premise of *Zebra* plausibly does depend on some background assumptions (2012, 277).

How close to either the middle or one of the peripheries of our spectrum a particular case belongs will depend on how clear the intuition of transmission failure is, how controversial this intuition is, and so on. *Toadstool* and *Twins* represent the most clear-cut cases at either end of the spectrum, with *Moore* representing the most controversial case in the middle. But there will be plenty of space in between these positions for varying degrees of controversy and strength of intuition.

With the landscape of the debate mapped in this way, one point that ought to be obvious is the following. Whether a particular theory of warrant is able to accommodate the phenomenon of transmission failure does not depend on whether it makes a particular prediction about the controversial Moorean case. Rather, the measure for whether a theory is able to accommodate the phenomenon ought to be whether it gets the peripheral cases right. And yet, this point often gets overlooked in debates about transmission failure, where very often the standard by which a theory is said to be able to accommodate transmission failure at all is whether it predicts transmission failure in *Moore*. Thus, in order to rebut the claims, quoted at the end of the previous section, that the externalist simply cannot say anything satisfying about the phenomenon of transmission failure, it will suffice to show that the externalist can make the correct prediction of transmission failure in the peripheral cases—i.e. in *Twins* and *Coins.* There will still be a question about what to say about the more controversial cases such as *Moore*, and the externalist who wants to deny transmission failure in controversial cases will still have to say what is the source of the unease about the kind of reasoning such cases represent, but it will not be true that denying transmission failure in *Moore* amounts to an inability to accommodate the phenomenon of transmission failure in general.

[67] Or some equivalent proposition.

3.3. First Attempt at an Externalist Account (*Zalabardo*)

The majority of the discussion around transmission and its failure, has been had by defenders of various forms of internalism. This raises the question of what externalists ought to say about the phenomenon. One possibility is that cases of transmission failure arise only insofar as we have already bought into an internalist framework. If this is the case, to the extent that one might not be attracted to the internalist view of epistemology, one need not worry about this apparent phenomenon. However, this does not seem like a plausible story to tell in light of the above-mentioned 'clear-cut' cases of transmission failure. The existence of clear-cut cases suggests there is a genuine phenomenon of transmission failure that is independent of our pet theories of knowledge or justification. The externalist therefore does need to be able to offer a diagnosis of transmission failure.

Recall Zalabardo's dilemma. Either we account for transmission failure by accepting (what he takes to be 'implausible') internalist constraints on warrant or we allow for unrestricted transmission even in cases where this clashes with our intuitions. Zalabardo finds both options unattractive, opting instead to try to account for intuitions of non-transmission without endorsing internalism. The last sentence could continue, 'nor indeed *any* constraints'. For Zalabardo's strategy is to argue that an unrestricted principle of transmission can accommodate our intuitions in all cases. Here is the principle he defends:

> *Transmission**
> If P is a proposition that S believes and for which S has warrant (and Q is a proposition for which S doesn't have warrant), then by recognizing the validity of the inference from P to Q, S will acquire warrant for Q

Zalabardo argues that Transmission* is defensible as it stands. The basic thought behind this is that in all problematic cases it is going to be impossible to satisfy the antecedent of Transmission*; impossible for S to have warrant for P but not for Q. In order to vindicate this thought, Zalabardo appeals to the following formulation of closure:

> *Closure**
> If P entails Q, and S has warrant for P and for the proposition that P entails Q, then S has warrant for Q.[68]

[68] This formulation comes from Zalabardo (2012, 312).

Zalabardo also appeals to the less familiar notion of epistemic transparency:

> *Transparency*
> P transparently entails Q if and only if P entails Q and anyone who possesses the concepts that figure in P and in Q has warrant for the entailment.[69]

The thought is that in alleged cases of transmission failure, the inferences in question satisfy *Transparency*, meaning that a subject who grasps the concepts at issue has warrant for the entailment regardless of whether they have carried out the inference. In combination with Zalabardo's formulation of closure (*Closure**), this entails that it will be impossible to satisfy the antecedent of *Transmission** and thus, actually carrying out the inference can do nothing to enhance the epistemic standing of the conclusion.

Here is how this works in the case of *Moore*. First assume that Transparency holds for *Moore* such that our subject S possesses the concepts that figure in (P) *here is a hand* and (Q) *there is an external world*, and thereby has warrant for the entailment. Simply by virtue of grasping the concept of a hand and of an external world, our subject S has warrant to believe that *P entails Q*. It now follows from *Closure** that it is impossible to have warrant for *P* without also having warrant for *Q*, so it is impossible to satisfy the antecedent of *Transmission**. Whenever one has warrant for *P,* one is guaranteed to have warrant for *Q*. Thus, there is no way to acquire warrant for the latter by means of carrying out the inference.

The first thing to note about this account is that, unlike orthodox accounts of transmission failure, it does not attempt to diagnose the issue by means of any kind of epistemic circularity. Whereas Wright's conservative account holds that transmission fails because warrant for Q is part of what gives one warrant for P in the first place, the thought behind Zalabardo's account is rather the idea that warrant for P and for Q are acquired in the same stroke, so the inference does nothing to improve one's epistemic standing towards Q.

An initial worry about this approach is that this notion of *epistemic transparency*, while plausible in general, seems problematic in the current context. Plausibly, some basic entailments are such that to understand the concepts involved is to have warrant for the entailment. Perhaps it is sufficient to possess the concepts that figure into *there is one apple and two oranges* and *there are more oranges than apples* in order to have warrant for the

[69] This formulation comes from Zalabardo (*ibid*.).

entailment, *if there is one apple and two oranges, then there are more oranges than apples*. But it is not obvious that the inference in *Moore* is like this, and almost certainly not the case that all cases of transmission failure are like this.

Consider the fact that there are alternatives to an external, material world ontology, such as for example the idealist mind-dependent picture of reality. If the inference in Moore is indeed epistemically transparent then idealists are conceptually confused. It is worth noting that even Moore himself argues at great length in his original 1939 article for the entailment from 'here is a hand' to 'there is an external world'. Such sustained argumentation conducted across several pages surely casts doubt on the possibility that Moore's inference is epistemically transparent. Idealism may ultimately be a false ontology but even so, it is not trivially false, and thus the inference in Moore cannot be transparent in the way that Zalabardo suggests it is.

Leaving issues around *Moore* to one side, an even bigger worry for Zalabardo's account is that it appears to run into trouble in meeting even the minimal desideratum outlined above. That is, it appears unable even to capture our intuitions in the clear-cut cases. The problem is that small changes to a case that affect whether we view it as transmissive or non-transmissive do not seem to affect whether we view the inference in question as epistemically transparent. Take our paradigm non-transmissive case. In *Twins* it is part of the background context that Sarah and Jessica are very hard to tell apart based on appearance. This feature of the case is essential in generating the intuition of transmission failure. Remove it—say for instance by specifying that Sarah and Jessica are not twins but merely sisters, who are in fact easily distinguished by their outward appearance—and the intuition goes away. However, it seems trivial that removing it makes no difference to whether the entailment is epistemically transparent: whether it is a conceptual truth that anyone who is Sarah is not Jessica is independent of whether Sarah and Jessica are visually indiscriminable. We can therefore imagine a second version of *Twins* in which there is no intuition of transmission failure because it is part of the background context that Sarah and Jessica are not twins and in fact have very different visual appearances. Call this second case *Sisters* If the entailment in *Twins* is epistemically transparent, then the entailment in *Sisters* must be epistemically transparent also. Therefore, (on Zalabardo's account) if there is transmission failure in *Twins* there is transmission failure in *Sisters* also. However, *Sisters* looks like as good a candidate as any for unimpeded transmission. This suggests that whether an inference is epistemically transparent has no bearing on whether it is non-transmissive.

Lastly, it has been pointed out elsewhere that Zalabardo's formulation of closure is implausible. It is implausible for the same reasons discussed back in Chapter I. when we rejected the more brute formulations of closure in favour of *Intuitive Closure*. Zalabardo's formulation is prone to the same kinds of trivial counterexamples in which one has yet to 'put two-and-two together', as it were, and recognise that Q follows from P. The principle should either be reformulated so as to weaken the consequent to *S is in a position to acquire warrant for P* (Wright 2012, 458). Or it should be reformulated more in line with *Intuitive Closure*[70] so as to make the antecedent more demanding, requiring that *S has carried out a competent deduction from p to q* (McGlynn 2013, Lockard 2016). Either way, the reformulated principles will not sustain Zalabardo's claim that it is impossible to satisfy the antecedent of *Transmission**. In both cases, even on the assumption that a subject has warrant for *P entails Q,* all that will follow from having warrant for P is that one is *in a position* to acquire warrant for Q. It will still be necessary to actually carry out the inference from P to Q in order to acquire warrant for Q, and thus still be possible to satisfy the antecedent of *Transmission**.

3.4. Second Attempt at an Externalist Account (*Smith*)

Martin Smith explores a different externalist-friendly strategy (2009). Smith exploits a distinction between what he calls *contributing* and *non-contributing* bases for belief. He identifies transmission failure with the failure to provide a basis for belief that contributes to the justification (reliability) of the belief. In order to see how this works, we first need to look at Smith's account of justification. In order for a belief to count as justified, Smith takes it to be necessary for the belief to be "held on a reliable or dependable basis" (2009, 176). Smith has a very specific way of understanding reliability, such that what it means for a proposition P to provide a reliable basis on which to believe another proposition Q is for P and Q to be "distributed in the right sort of way throughout certain important possible worlds" (*ibid*). Rather than the relationship being one of closeness to the actual world—as with more familiar modal conditions such as safety or sensitivity—Smith intends the relevant notion to be one of *normalcy*, relative to the actual world. A basis B "provides a reliable basis on which to judge that P just in case, in all those possible worlds in which B is true and which are as normal, from the perspective of the actual world, as the truth of B permits them to be, P is true too" (2009, *ibid*).[71] Although this notion of normalcy is a somewhat idiosyncratic way of explicating the

[70] Recall the formulation of this principle from chapter 1.

[71] This notion of normalcy allows for the non-factivity of justified belief. Sometimes abnormal conditions prevail. Consider, for example, a classic Gettier case in which I see what very clearly looks like a barn in a

notion of a reliable basis for belief, Smith takes it to be "perfectly in keeping with the ordinary use of the term 'reliability'" (2009, 177).

It follows that deductive inference can never take one from a reliable belief to an unreliable belief. Consider again the zebra case. My basis for belief that (P) *those animals are zebras* is (E) *the visual appearance of zebras in the zebra enclosure*. This basis is reliable on Smith's definition: in all of the most normal worlds in which E is true, P will be true also. Even if in the actual world the animals are in fact cleverly disguised mules, it will still remain true that zebra-ish appearances are a normally reliable indicator of the presence of zebras. Furthermore, given my reliable basis for P, I cannot fail to have a reliable basis for (Q) *those animals are not cleverly disguised mules*. This is, of course, because if all of the most normal worlds are P-worlds, then all of the most normal worlds are Q-worlds too (given that P entails Q). There won't be any normal worlds in which *those animals look like zebras* is true while *those animals are not cleverly disguised mules* is false. Thus, if E is a reliable basis on which to believe that P, then E is a reliable basis on which to believe that Q.

Smith distinguishes between two ways that a belief can be reliably based. A basis B provides a *contributing* basis for belief that P so long as, in addition to all the most normal B worlds being P worlds, all the most normal ~P worlds are ~B worlds (2009, 179). It follows that the appearance as of zebras in the enclosure provides a contributing basis for belief that the animals are zebras, given that if there were not zebras in the enclosure, one would not normally expect to see things that looked very much like zebras. In contrast, a basis B provides a mere *non-contributing* basis for belief that P so long as, in addition to all the most normal B worlds being P worlds, some of the most normal ~P worlds are still B worlds (*ibid*.). It follows that the appearance as of zebras in the enclosure provides a non-contributing basis on which to believe that the animals are not cleverly disguised mules, given that if in fact there were cleverly disguised mules in the enclosure, we would of course find things that looked like zebras in the enclosure. When a basis for belief is reliable but non-contributing in this sense, the belief in question is reliable in virtue of its *content* (2009, 180). It is the fact that the truth of (Q) *those animals are not cleverly disguised mules* is partly constitutive of what it means for conditions to be normal that such propositions are, in a sense, normic necessities (true in all the most normal worlds). This, according to Smith's account, is what it is for an inference to fail to *transmit* justification. An inference whose premises provide a mere non-contributing

nearby field. In all the most normal worlds in which I clearly see what looks like a barn in a field, it will be true that there is a bard in a field, and thus my barn-belief is reliably based. And this will remain the case even if, in the actual world, I happen to be looking at a barn façade.

reliable basis for belief in its conclusion fails to enhance or transmit reliability, and since reliability is necessary for justification, this is sufficient for an inference to fail to transmit justification.

Smith's account clearly predicts transmission failure in *Moore*. Once again, the conclusion of *Moore* is going to be such that its truth is partly constitutive of what it means for conditions to be normal. From the perspective of the actual world, the world in which there is in fact no mind-independent reality is very abnormal indeed. So firstly, Smith's account allows that Moore's argument obeys closure. The evidence (E) *it appears as though I have two hands* provides a reliable basis on which to believe (P) *I have two hands* because all of the most normal worlds in which E is true P is true too. This basis is of the contributing variety, since some of the most normal worlds in which one does not have hands are worlds in which it does not appear as though one has hands. Furthermore, with the deductive inference from P to (Q) *there is an external world*, this in turn provides a reliable basis on which to believe that Q, since again, the nature of deduction means that all P worlds will be Q worlds also. However, this basis is of the non-contributing variety given that the truth of Q is partly constitutive of what it means for conditions to be normal and thus Q is reliable in virtue of its content. If Q were false, there would of course still be hand-ish appearances. As Smith puts it:

> The appearance of hands before my face will not make a world in which reality is mind dependent any less normal than it would otherwise be. That is, the most normal worlds in which reality is mind dependent and there appear to be hands before my face will simply be amongst the most normal worlds in which reality is mind dependent. My account predicts, then, that Moore's inference fails to transmit justification from premises to conclusion. One cannot augment the reliability of a belief in the existence of the external world by holding one's hands before one's face. (2009, 175-176).

___

There are four issues for Smith's account that I want to draw attention to. It appears as though Smith's account runs into trouble with the *Twins* case above. Recall the conclusion of this case. For Smith's account to get the result that transmission fails in *Twins*, it would need to be the case that the proposition *that person is not Jessica* enjoys epistemic support in virtue of its content alone, rather than any support it gets by virtue of the inference. In particular, it would need to be the case that the truth of this proposition is partly constitutive of what it means for conditions to be normal; that the falsity of this proposition would represent very abnormal conditions. However, in thinking about the proposition in question—*this person is not*

Jessica—that does not seem to be the case. Recall the details of the example. Sarah and Jessica are twins whom we find ourselves unable to tell apart. We justifiably believe that *this person is Sarah* partly in virtue of the visual experience and partly in virtue of some sort of background justification for believing that Sarah's twin, Jessica, is not in the vicinity—above we assumed that someone had informed us that Jessica is out of town on a business trip. But notice that it does not seem right to say the truth of the conclusion—*this person is not Jessica*—is partly constitutive of what it means for conditions to be normal. Although we justifiably believe that Jessica is out of town, it would not be out of the ordinary if this belief turned out to be false; if, for example, Jessica missed her flight, or returned home early for some reason. This would be a very ordinary way in which the belief in question could be false—notwithstanding the reliable basis we have for thinking it is true. This means that the case does not fit into Smith's normic account of when justification fails to transmit from premise to conclusion. Perhaps this is not ultimately all that problematic for the account if a distinction can be drawn between this case and the kinds of cases that the account is directed at. Even so, Smith would then owe us a principled account of the grounds on which we ought to draw such a distinction.

The next issue that Smith's account of transmission failure faces concerns inferences whose conclusions are necessarily true. A belief in a necessary truth cannot fail to be reliably based, given the above account. Where I believe some necessary truth (P) on some basis (B), B will count as a reliable basis to believe that P, purely in virtue of the fact that there are no worlds—normal or otherwise—in which P is false. As Smith flags, the distinction between contributing and non-contributing bases for belief breaks down when it comes to beliefs in necessary truths (2009, 171). This is because, while any basis cannot fail to be a reliable basis on which to believe a necessary truth (given Smith's definitions), we cannot run the counterfactual test for whether that basis is of the contributing or non-contributing variety (because we cannot counterfactually suppose the falsity of a necessary truth). Nonetheless, Smith claims that "If P is a necessary truth then any basis will qualify as a contributing reliable basis, as defined here, upon which to believe that P" (2009, 179-180). It is not entirely clear why the account must force this result, and I think a case can be made in the opposite direction. For consider, the key idea behind non-contributing bases for belief is that the beliefs in question are reliably based purely in virtue of their content. The propositions that Smith has in mind are those that are true in all of the most normal worlds—these propositions are, in a sense, *normic necessities*. And the reason why certain bases fail to contribute to the reliability of normic necessities is that these propositions are already supremely reliable in virtue of their content alone. But then necessary truths are also true in all of the most normal worlds, indeed true in

all worlds. A natural conclusion to draw from this is that if we are to view *normic* necessities as having this distinctive property of being reliable in virtue of their content, then we ought to view bona fide necessary truths as reliable in virtue of their content too. This is not the conclusion that Smith himself draws, but it is not obviously the wrong conclusion to draw. But then suppose that this is the case, it now looks like inferences whose conclusions are necessarily true must fail to transmit justification. Suppose that using an electronic calculator I calculate that (P) *the square root of 289 is 13*, and then infer that (Q) *the square root of 289 is less than 15*. Presumably, we ought to count this inference as transmissive of justification. However, given that Q is a necessary truth, it will be maximally reliable purely in virtue of its content, and as such the inference fails to enhance Q's reliability in any way. If the diagnosis of *Moore* is that the inference fails to make its conclusion any more reliable than it already is in virtue of the fact that there are no normal worlds in which it is false, then ought we not say—with perhaps even more assuredness—that the inference here similarly fails in virtue of the fact that there are no worlds simpliciter in which it is false? I do not see how the account can avoid this result. But suppose we reject this idea that we must view these inferences as instances of *non-contributing* bases for belief and agree with Smith that they are, in fact, instances of *contributing* bases for belief. Well then, we get another counterintuitive result, namely that it is impossible for an inference to fail to enhance the reliability of its conclusion if that conclusion is necessarily true. Consider the inference *the square root of 289 is 13, therefore the square root of 289 is 13*. Again, this is simply another instance of a question-begging inference which is one of our paradigm cases of transmission failure. Such inferences can have no hope of providing any kind of epistemic boost, as it were. That is, we do not end up in a better place than where we started out by simply recognising that a proposition entails itself. And yet, if we suppose that inferences whose conclusions are necessarily true cannot fail to provide *contributing* bases for belief, we get the result that these inferences do indeed enhance the reliability of their conclusions. But it is very difficult to see what could be being enhanced here. The only solution seems to be to exclude from the present account inferences whose conclusions are necessarily true. One would then need a principled way to make such a division, and a separate account for how to deal with these inferences.

The third problem for Smith's account concerns how we stipulate inferential beliefs are based. Consider again the counterfactual test for a non-contributing basis for belief: "B is a non-contributing reliable basis upon which to believe that P just in case, in some of the most normal worlds in which P is false, B is still true" (Smith, 179). And recall how this works in the zebra case: in some of the most normal worlds in which it is false that (Q) *those animals*

are not cleverly disguised mules it is true that (E) *those animals look like zebras*. So if I base my belief that Q on E then that basis for belief counts as non-contributing according to the counterfactual test. However, suppose we tell a subtly different story according to which E provides a reliable basis on which to believe that P, and P in turn provides a reliable basis on which to believe that Q. This alternative picture is both plausible and theoretically significant. It is plausible because one natural way to think about inferentially based beliefs is that, by inferring Q from P, I thereby base my belief that Q on P, rather than on whatever evidence I take myself to have for P. According to this picture, once I have based my belief that P on whatever defeasible evidence I have for P, I am now free to reason with P and base further beliefs on P itself rather than on the original evidence for P. Assume this picture is correct such that I base my belief that (Q) *those animals are not cleverly disguised mules* on my belief that (P) *those animals are zebras*, rather than on whatever visual evidence I have for P. There are now significant theoretical consequences for Smith's account. My basis for belief that (Q) *those animals are not cleverly disguised mules* now fails to meet the definition of non-contributing. Rather, it meets the definition of a *contributing* reliable basis for belief: "B is a contributing reliable basis upon which to believe that P just in case […] in all the most normal worlds in which P is false, B is false too" (*ibid*). In all the most similar worlds in which it is false that those animals are not cleverly disguised mules, it is false that those animals are zebras. Thus, if I base my belief that those animals are not cleverly disguised mules on the proposition *those animals are zebras*, then according to Smith's definition this is a contributing reliable basis for belief, and the inference is transmissive after all. The upshot is that whether the account predicts transmission failure in cases that exhibit the familiar tripartite structure (E-P-Q) seems to depend on an unargued for assumption about how beliefs in Q-type propositions are based.

We might wonder whether there is a way to avoid this last problem by insisting that where one has a justified belief that P based on E, and then goes on to infer Q from P, one's belief that Q is ultimately always based on E. I suspect that perusing this strategy as a way to neutralise the above problem is going to seem rather ad hoc. Simply insisting that inferentially based beliefs are always ultimately based on the initial evidential basis one has for the premise belief seems like a not particularly intuitive psychology of belief formation. For example, consider cases in which one has a justified belief that P even though one has lost one's initial evidential basis for P. I take myself to have a justified belief—indeed, to *know*—that *Canberra is the capital city of Australia*. I am, however, entirely ignorant of the basis on which I formed this belief. Nonetheless, I take myself to be in a position to infer that *the capital city of Australia is not Adelaide*, and for this newly inferred belief to be justified. Insisting that this new belief

is ultimately based on my initial evidential basis for the former belief does not seem like an independently plausible idea.

The fourth and final issue that I want to raise concerns the counterintuitive result that a belief which is reliable purely in virtue of its content seems thereby to be immune from unreliable basing. For example, Smith's account tells us that when Moore bases his belief that (Q) *there is an external world* on the inference, the reliability of his belief is accounted for by the belief's content alone. Crucially, the inference contributes nothing to the reliability of this belief given that it has the particular content it does. But the problem then is that it looks like anything will count as a reliable basis on which to believe such a proposition. Suppose that I happen to believe that *there is an external world* because a ghostly apparition came to me in a dream and declared it to be so. I take it that, whatever support this proposition may enjoy as a matter of being normically supported in virtue of its content, if I happen to believe it on such a flimsy basis, my belief will not be reliably based. Or suppose that I am considering which of two hypotheses to believe: (A) *that reality is mind-independent* or (B) *that reality is mind-dependent*. I decide that I'm going to flip a coin to settle the matter and believe either A or B based on the outcome of the toss. I take it as given that this is an unreliable method of belief formation—indeed, it is the *paradigm* case that philosophers often appeal to in example. Even if I arrive at a true belief, my belief will not be reliably based. And it will not be reliably based even though my true belief is true in all of the most normal worlds to the actual world. Smith's account of beliefs that are reliable in virtue of their content alone seems insensitive to the difference between cases like this and cases where the belief is based on, for example, the actual reasoning at issue in Moore's inference.

We have looked at four problems for Smith's reliabilist account of transmission failure. In the first two problems it looks like the account mistakenly treats some inferences as transmissive that it intuitively ought not, and likewise counts some inferences as non-transmissive that intuitively it ought not. The second two problems concern the way in which inferred beliefs are based. The account seems to presuppose a substantive conception of inferentially based beliefs which is not obviously correct, and seems counterintuitively to count paradigmatically unreliable bases for belief as, in fact, reliable. I conceded that there are perhaps ways in which the account can be modified to avoid at least the first two problems. I am more sceptical that the account can be modified to avoid all four problems.

3.5. The Simple Reliabilist Account of Transmission Failure

As indicated in the passages quoted earlier on in section 3.1, Wright is sceptical about the hopes of providing a satisfying account of transmission failure in terms of reliability or safety. Wright justifies his scepticism on the following grounds. Beginning with reliabilism, he argues that given that deductive inference is a paradigm reliable process, a deductive inference from a reliably formed belief cannot fail to inherit the entailing belief's reliability (455-56). However, he foresees a possible response the reliabilist could make to this argument. If the reliable method used already somehow involves a belief in the argument's conclusion, then rehearsing the argument won't be a way of *acquiring* a reliably formed belief in the conclusion. He goes on to reject this line of response: "This kind of play is no help in the crucial range of cases where the belief-forming method that leads to acceptance of the premises consists simply in the exercise of some non-inferential cognitive faculty—say, perception" (456). An interesting thing about these remarks is that the cases that Wright appeals to in illustrating the phenomenon of transmission failure are not exclusively ones that involve non-inferential beliefs. Indeed, the very case identified in section 3.2. as the crucial case for assessing whether a candidate theory can account for the phenomenon of transmission failure involves a premise belief that, on the most plausible interpretation of the situation, is formed inferentially.

Recall the premise of *Twins*: *this person is Sarah*. The most plausible reading of this case is one in which this belief is formed inferentially. Suppose otherwise: that this belief is formed non-inferentially via perception. One sees a person who looks like Sarah and immediately forms the corresponding belief. Such a belief seems like it is not formed on a reliable basis. After all, if one goes around believing that *this person is Sarah* regardless of whether the person one is looking at is Sarah or Sarah's twin Jessica, then this method will often lead to false belief. In order for the belief to be reliably based, the basis for belief needs to include the background information that the individual in question is not Jessica. As Wright himself puts it in his discussion of this case: "…you—though perhaps not someone who can distinguish the twins purely visually—will need [Q] already in place as collateral information before you can reasonably take E as a warrant for P" (2003, 60). So it looks like belief in the premise of *Twins*, on the assumption that we cannot tell the twins apart based on vision alone, is based on the visual evidence provided by E together with the background information that figures in the conclusion. So it is false that in *Twins* the belief is formed non-inferentially. If I am right in arguing that *Twins* is a crucial case that any theory of transmission failure needs to get right, then it looks like the kinds of response (discussed in the paragraph prior to this one) that Wright foresees then immediately rejects are, in fact, perfectly sustainable in at least this

type of case. A reliabilist account of justification can explain transmission failure in terms of an inferential basis for belief in the premise that already includes justification for the conclusion.

One term that has been used to describe the kind of epistemic deficiency that question-begging arguments such as the above exhibit is *epistemic dependence*. This term is used by Pryor to capture a number of different ways that arguments exhibit various interesting properties (2004, 357-362). Some of these properties are such that they give rise to transmission failure, while others seem less epistemically objectionable. The suggestion I am going to make here is that some of the particular forms of epistemic dependence that Pryor describes are available to the externalist to appeal to in order to account for transmission failure. Before getting to that, I first want to be very clear about the difference between the kind of view that I am going to propose and the kind of view of transmission that Pryor endorses. In a certain sense, the view that I am going to propose is very much like Pryor's in that it is going to treat some of what I above labelled the 'controversial' cases as cases of successful transmission. However, one very important difference is that Pryor's account of justification is unequivocally internalist. For example, with respect to the Moorean inference, which Pryor takes to be a case of case of successful transmission, he claims that the justification for the premise—*here is a hand*—is accounted for in terms of the phenomenology of the experience as of seeing a hand (2000, 536). The fact that one has a perceptual experience of a certain sort is enough on its own to provide one with justification to hold the corresponding belief, albeit a justification that is defeasible. The view that I am going to advance rejects this phenomenological conception of justification in favour of a more externalist notion.

Pryor discusses a number of different ways in which the premises of an inference can depend on the conclusion. When the premises of an inference depend on the conclusion in an epistemically objectionable way, Pryor describes the kind of dependence involved as an "epistemic vice" (2004, 359). Not all forms of epistemic dependence are vices. Three kinds of dependence are going to be most relevant for the current discussion. The first type of dependence will be familiar from the above section on Zalabardo's diagnosis of *Moore*. Some arguments are such that, in order to have justification for the premises, one needs to have justification for the conclusion also.[72] As Pryor notes, this kind of dependence is going to be harmless in cases where the inference is "*so obvious* that understanding the premise well enough to be justified in believing it *requires* you to take any justification for the premise to

[72] This is what Pryor labels "Type 3" dependence (2004, 359).

also justify you in believing the conclusion" (2004, 359). The idea that Pryor is getting at here is very similar to what Zalabardo labels transparency. And again, there are plausibly going to be some inferences that exhibit this type of epistemic dependence. It is not clear that we ought to view this kind of dependence as epistemically objectionable in the way that ordinary cases of transmission failure are epistemically objectionable. On the one hand, this kind of structure gives rise to transmission failure for the reason that it rules out acquiring justification for the conclusion—or acquiring a justified belief in the conclusion—by virtue of rehearsing the inference. Inferences that exhibit this type of dependence are such that rehearsing the inference is not a way to improve one's epistemic standing towards the conclusion. On the other hand, the reason for this is not that the inference exhibits any kind of epistemic circularity in the way that genuinely problematic cases of transmission failure do. Rather, justification for the conclusion comes synchronously with justification for the premise. It will be helpful to have a label to refer back to this type of dependence so, for want of a better term, I will call this *transparent dependence*.

The next type of epistemic dependence that Pryor discusses involves inferences in which evidence against the conclusion would serve to undermine the justification for the premise (to at least some degree). Pryor labels this type 4 dependence but I am instead going to call it *negative dependence*. Many of the arguments that we have been discussing such as *Moore* and *Zebra* exhibit *negative dependence*. If I acquire some evidence that, in fact, the animals in the pen *are* cleverly disguised mules, then this will undermine the justification for the belief that the animals are zebras based on the visual evidence as of their appearing to look like zebras. Likewise, if we can imagine a scenario in which one could acquire some evidence against the proposition *the external world exists* then that would serve to undermine the perceptual evidence for the belief that one has hands.

Negative dependence needs to be distinguished from a trivial property that all deductive inferences have. Any deductive inference is such that evidence against the conclusion will count as evidence against the premise, and thus may serve to defeat justification for the premise. But, argues Pryor, evidence against the conclusion may serve to defeat the justification one has for the premise without *undermining* that justification. The distinction is a subtle one. To illustrate the point, Pryor considers a case in which I believe that (P) *I ingested no memory altering drugs recently* (2004, 373). Suppose that I believe this proposition on the grounds of either (A) *testimony from a reliable source* or (B) *my own memory*. Having some reason to doubt P would defeat the evidence provided by either A or B. But only with respect to B would the evidence against P *undermine* B. If I had evidence for thinking that I *had* taken

memory altering drugs, this need not undermine the testimonial evidence, it might simply indicate that the generally reliable source was wrong on this occasion. But evidence that I had taken the memory-altering drugs would undermine the memorial evidence represented by B. Where a deductive inference exhibits mere negative dependence, this is where evidence against the conclusion would serve to undermine the evidential basis for the conclusion.

Contrast this to a different type of dependence that Pryor highlights in which having justification for the conclusion is among the conditions that give you the justification for the premise (2004, 359). I will label this *positive dependence*. As Pryor recognises, many inferences that exhibit negative dependence will also exhibit positive dependence. To view an inference as exhibiting positive dependence is, in Pryor's terminology as mentioned above, to treat the inference conservatively. And since it is the conservative diagnosis that gives rise to transmission failure, any inference that exhibits positive dependence therefore exhibits transmission failure. For example, it is trivial that the simple question-begging inferences such as *P, therefore P* exhibit positive dependence—that is, in order to have justification for the premise (P) one needs justification for the conclusion (P). But while this kind of dependence is a bad thing, it is unclear why negative dependence would be a bad thing.

Some inferences might be such that they exhibit (merely) negative dependence but not positive dependence. Pryor argues that there ought to at least be some inferences that have this structure. For example, certain introspective beliefs such as *I am feeling cold* may be such that justification for them can be defeated by evidence indicating one is committing a priming mistake without that implying that one needs prior justification to believe that one is not committing a priming mistake. A priming mistake is where one is 'primed' to expect a certain sensation such as 'feeling cold' which results in one mistakenly judging actual sensations of heat for sensations of cold. If such cases are possible, then negative dependence does not entail positive dependence and the space is opened up to argue that what we find in cases such as *Moore* is mere negative dependence. What such a diagnosis allows us to do is reject the claim that *Moore* exhibits transmission failure, while opening up the space to explain what else could be wrong with the inference.

The idea of an inference exhibiting mere negative dependence is closely connected to the kind of response to Wright given by Martin Davies (2004). Davies claims that Wright is overly concessive to the sceptic. From the claim that doubt about the Q-type proposition would serve to defeat any justification one would have for the P-type proposition, it need not follow that one needs positive justification to accept the Q-type proposition. Perhaps absence of reasons to doubt is sufficient. This is a possibility that Davies suggests Wright overlooks, and

it is one Davies exploits in developing his notion of a *negative entitlement* (2004, 226). The views advanced below for how the externalist ought to deal with transmission failure will resemble Davies' account in certain important respects, although Davies account does not draw this further distinction along the lines of what I am calling *transparent dependence*.

Nothing in this distinction between negative, positive and transparent dependence presumes anything about internalist or externalist accounts of justification. Pryor appeals to this distinction to account for a difference between arguments that exhibit transmission failure and arguments that may superficially seem like they do, which he then combines with his preferred internalist account of justification. The very minimal point that I wish to make here is simply that the distinction between these three types of dependence is open to the externalist to appeal to as well. We have already seen how a simple reliabilist account of justification might appeal to the idea of positive dependence to explain how justification fails to transmit in cases such as *Twins*. Wright himself acknowledges that this kind of move will be open to the reliabilist in cases where the reliable basis for belief in the premise already involves a reliable belief in the conclusion (2012, 456). However, as Wright goes on to point out, this kind of strategy will not work in cases where the premise belief is formed non-inferentially. If the premise belief is formed immediately via, say, perception, then no prior belief in the conclusion will be involved. To the extent that there are clear-cut cases of transmission failure that do not involve inferentially based premise beliefs, this looks like an uncomfortable result for the reliabilist who wants to account for transmission failure. I will next try to show how the reliabilist can avoid this result in at least certain cases by appeal to this notion of transparent dependency.

Recall again the coins case from section 3.1. I count that there are seven coins on the table, form the corresponding belief that there are seven coins on the table, then infer that I have not miscounted. What can our basic reliabilist account of justification say about such an inference? The first thing to say is that the inference clearly exhibits negative dependence. Any evidence that indicates that I in fact miscounted will serve to undermine the justification I have for believing based on my count. I will not be able to rely on the evidence provided by the count for dismissing the evidence that suggests I miscounted. So, the inference from *I counted seven coins on the table, therefore there are seven coins on the table* to *therefore I did not miscount*, does exhibit negative dependence. Does the inference also exhibit positive dependence? Certain accounts of justification are going to answer this question in the positive. The kind of conservative view that Wright defends will say that yes, that one did not miscount is an authenticity condition for the project of counting the coins, and therefore one is committed

to taking this claim on trust. On the other hand, it looks like our basic reliabilist picture is going to answer the question in the negative. All that one needs in order to justifiably believe that there are seven coins on the table is that one form this belief in a reliable way. This justification can be easily undermined—in Pryor's sense—by some evidence to believe that one did in fact miscount. But this is simply to point out that the inference exhibits negative dependence. The reliabilist account of justification therefore seems like it is committed to rejecting the thought that this inference exhibits positive dependence, and therefore rejecting the thought that it exhibits transmission failure. Is this a bad result for the reliabilist? It is a bad result if there are independent reasons for thinking that inferences of this type do exhibit, not merely negative, but positive dependence. But in the absence of such reasons, what a basic reliabilist account might say about such inferences is that they do not exhibit transmission failure. Any intuition to the contrary, unless backed up by an argument as to why we ought to view such inferences as exhibiting positive dependence, ultimately seems to amount to what Brian Weatherson calls a "whiff of circularity" (2019, 147). And it is then open to the externalist to diagnose this whiff along the lines of mere negative dependence.

In addition to what I have been calling negative and positive dependence, I also discussed a type of dependence that Pryor highlights and which, drawing inspiration from Zalabardo, I labelled *transparent dependence*. This is the idea that justification for the premise might depend on justification for the conclusion if, whenever one has justification for the premise, one also has justification for the conclusion. I remarked above that it is unclear whether we ought to view this kind of dependence—if we can even call it that—as epistemically objectionable. While I argued earlier on that Zalabardo is mistaken in diagnosing *Moore* as exhibiting this phenomenon, we might be tempted to say that our coins case does exhibit transparent dependence. The thought here would be that simply by virtue of counting that there are seven coins on the table one thereby acquires justification to believe that one has not miscounted. This idea seems particularly plausible once we recognise that such justification would be defeasible, so that as long as one has no reasons or evidence to think that one has miscounted, in counting one thereby acquires justification to believe that one has not miscounted. Where this view would differ from, say, Wright's entitlement account is that it would deny that such justification or warrant to believe that one has not miscounted is in place by *default*. Rather, one *earns* the justification by virtue of having counted the coins in a certain way. The reliabilist will, naturally, hold that so long as one performs the count in a reliable manner and that there are no defeaters for the belief that one has not miscounted, one *thereby* acquires justification to believe that one has not miscounted. If this idea makes sense, then we

have a reliabilist story of where the source of the intuition that transmission fails in this case comes from. Although the inference does not exhibit positive dependence, it does exhibit transparent dependence. The justification to believe that there are seven coins on the table and the justification to believe that one has not miscounted are both acquired in the same stroke, as it were. This means that one cannot acquire the latter by means of rehearsing the inference. Once again, it is not clear whether we ought to view this kind of dependence as an epistemic vice given that it does not presume there to be any kind of circularity in play.

As we have seen, the externalist is able to accommodate the phenomenon of transmission failure in at least the most clear-cut cases. However, the question remains of what to say about the more controversial cases such as *Moore*. The simple reliabilist account of justification that I have been appealing to in order to capture intuitions in the twins and coins cases looks very much like it will predict successful transmission in *Moore*. After all, one might reliably form the belief that (P) *here is a hand* without having formed any beliefs about the external world, so what is to prevent one from then recognising the entailment and reliably forming a belief in (Q) *there is an external world*. The view that I want to endorse says there need be nothing preventing us from reasoning in this way. However, we began chapter one by introducing the Moorean inference and noting that many find it to be an impotent response to the sceptic. So we need to say something about the source of the intuition that there is something fishy about the inference. The first thing to say about the Moorean inference is that it clearly exhibits negative dependence. This is part of the story that Pryor tells about this case and there would seem to be no reason why the simple reliabilist picture under consideration cannot make use of the same story. Evidence against the conclusion would undermine the evidence of the senses for forming empirical beliefs about the existence of objects.

In addition to account for intuitions regarding Moore along the lines of negative dependence, notice that what is interesting about inferences that exhibit negative dependence is that they will exhibit a further property as well. If an inference is such that doubt about its conclusion will serve to defeat any justification for the premise, then, as Pryor points out, "if [a] subject *just happens to doubt* the conclusion, without having any reason for doing so, her doubt will rationally obstruct her from believing the argument's premises (2004, 335-336). This idea is echoed by Wright when he introduces the notion of an authenticity condition and the general tripartite (E-P-Q) structure of cases such as *Moore*:

> […] doubt about [Q] is rationally precluded on the part of one who proposes to take the evidence for the [P]-proposition depicted by [E] as sufficient. Absent other relevant information, any doubt about the [Q]-proposition must tend, in a rational subject, to undermine the force of the evidence described in [E] for the [P]-proposition. Such a doubt may, as in the case of Moore, Zebra [and so on] involve doubt about the [P]-proposition too. But in all cases, it must involve doubt about the significance of the evidence depicted in [E]. One who doubts the [Q]-proposition could not rationally move to belief in the [P]-proposition just on the basis of evidence [E].[73]

Curiously, what Wright is describing here is not any kind of conservatism (though he does develop his conservatism later on in the same article). This passage is something that liberals can get on board with. A similar notion lies behind his notion of a presupposition for a cognitive project, defined in his earlier article on entitlement:

> P is a *presupposition* of a particular cognitive project if to doubt P (in advance) would rationally commit one to doubting the significance or competence of the project.[74]

And once again, liberalism is perfectly compatible with the existence of presuppositions in this sense. One effect of merely doubting the existence of an external, mind-independent world will be that one can no longer rationally believe that one has hands on the basis of the visual evidence. This is obviously different to the claim that if one has evidence against the external world claim that one's justification for the belief that one has hands gets undermined. The claim here is purely about rational commitments, and one may be rationally committed to doubting that one has hands even when one has sufficient justification to believe that one has hands. This is a doubt that, all things considered, one ought not have because one's doubt in the existence of the external world is itself not supported by evidence. But given that one does in fact doubt the existence of the external world, this rationally commits one to doubting both that P and that E justifies P. Nothing I am saying here has not already been said elsewhere. Both Pryor (2004, 2012) and Davies (2004) diagnose the Moorean inference as suffering from certain *dialectical* deficiencies; the inference cannot be used to overcome doubt about its conclusion and is thus no good in a dialectical standoff with a sceptical opponent. The only

[73] Wright (2014, 214).
[74] Wright (2004, 191).

(admittedly, very minimal) claim that I make is that this diagnosis, in terms of doubt and rational commitments, is open to the externalist who rejects other aspects of, for example, Pryor's epistemology such as the claim that perceptual justification is a matter of having a certain kind of phenomenological experience.

3.6. Prospects for a Safety-Based Account of Transmission Failure

Earlier in chapter one we distinguished between three kinds of questions that we might want to ask about closure and transmission, relating to propositional justification, doxastic justification and knowledge. In line with the vast majority of literature on transmission and transmission failure, most of our focus since then has been on justification, and in particular doxastic justification, and I have argued that a reliabilist account of doxastic justification can accommodate transmission failure in at least some of the most clear-cut cases. One might be tempted to think that this also gives us an answer to the question of whether there are counterexamples to transmission of knowledge, since justification is standardly taken to be necessary for knowledge.[75] There is, however, a different strategy we might take in order to try to approach this question of whether there is transmission failure of knowledge. Many philosophers find plausible the idea that a necessary condition on knowledge is some sort of modal condition such as safety. The very rough idea behind safety is that a belief is safe so long as it could not easily have been false. More precisely, a belief is safe so long as it is true the actual world as well as all or most nearby possible worlds too.[76] Thus, if we want to know whether knowledge violates transmission principles, one thing we can ask is whether safety violates them.

---

[75] The suggestion that we can infer things about whether knowledge transmits from conclusions about whether justification transmits given the necessity of justification for knowledge is reminiscent of an exchange between Ted Warfield (2004) and Anthony Brueckner (2004). Warfield challenges the claim that we can establish that knowledge is not closed by establishing that some necessary condition on knowledge is not closed. His point is that from the non-closure of some condition (C) that is necessary for knowledge, the non-closure of knowledge need not follow so long as there could be some further necessary condition (D) which ensures that a subject, S, satisfies C. If Warfield is correct, we might be inclined to think that similar points apply to transmission as well as closure. In what follows, I will attempt to draw some conclusions about whether knowledge transmits from observations about whether safety transmits. So it does look like Warfield's argument, if successful, would undermine what is to follow. However, I think that Brueckner makes a convincing case in response to Warfield (2004). Brueckner's point is effectively that, so long as we take the necessary conditions for knowledge to be logically independent of each other, there is no reason to think that satisfying one condition—D, say—guarantees that one satisfies another—C, say. Moreover, the discussion that follows is largely a response to Smith, who does also seem to work under the assumption that observations about whether the property of safety transmits tell us something about whether knowledge transmits. I thus inherit this assumption from Smith and refer the reader to the Warfield-Brueckner discussion to make up their own mind whether it is warranted.

[76] This formulation is overly simplistic and neglects to mention the importance of holding fixed the basis on which the belief is formed. It will, however, do for now and the importance of basis-relativity will be made explicit below.

Wright is sceptical about the hopes for a safety-based account of transmission failure: "the […] idea of a true belief that holds good in all reasonably close worlds in which it is formed in the way it is actually formed is obviously hereditary across valid inferences leading to beliefs in their conclusions" (456). These remarks are perhaps a little quick. What is *obvious* is that—on a certain conception of safety[77]—a proposition that is validly inferred from a safe proposition cannot fail to be safe itself. But this is simply to say that safety obeys closure, which is a truism. If P is true in all nearby worlds, then trivially all of P's deductive consequences will be true in all nearby worlds because, as a matter of logical necessity, deductive inference cannot introduce any new error possibilities.

It remains to be seen whether safety always transmits from premise to conclusion. Martin Smith has a safety-based account of transmission failure that is complimentary to his reliabilist account of transmission failure for justification discussed above. Once again, the very general idea driving Smith's approach is a distinction between contributing and non-contributing bases for belief. While deductive inference cannot take one from a safe belief to an unsafe belief, it can take one from a belief that is safe in virtue of its basis to a belief that is safe in virtue of its content, thereby providing a non-contributing basis for belief in the conclusion. And as with his reliabilist account of transmission failure above, Smith argues that cases of transmission failure are those for which the inference provides a non-contributing safe basis for belief (2009, 170).

To see how Smith gets this result, we should first get a handle on what it is for a belief to be safe. Taking his cue from Ernest Sosa, Smith defines safety along the following lines: a basis B "provides a safe basis upon which to judge that P just in case, in all those possible worlds in which B is true and which are very similar to or closely resemble the actual world, P is true too" (2009, 168). For example, the appearance as of animals that look like zebras provides a safe basis to believe that they are zebras. Suppose that I see animals in the zebra enclosure that (B) look like zebras. In all nearby worlds it will be true that (P) the animals are zebras, and thus B provides a safe basis on which to believe that P. This also provides a safe basis on which to believe that (Q) the animals are not cleverly disguised mules. In all the nearby worlds in which the animals (B) look like zebras, it will be true that (Q) they are not cleverly disguised mules. Although there will be some worlds in which B is true and Q is false, these

[77] I have in mind a very minimal conception of safety that should be compatible with a variety of accounts such as Williamson (2000, 147), Pritchard (2007a, 2008), Sosa (1999a, 1999b). For a helpful discussion of the differences between these accounts see Rabinowitz (2011).

worlds are sufficiently far-off from the actual world so as to render them irrelevant for the purposes of evaluating safety.

As with his account of reliability, there are two ways that Smith thinks a proposition B can provide a safe basis for belief in another proposition P. A belief that P can be safe in virtue of the modal profile of B and P. In such a case, B provides a *contributing* safe basis for belief that P. Alternatively a belief that P can be safe regardless of its basis, in virtue of its content alone. In such a case, B provides a non-contributing basis for belief that P. Here are Smith's definitions of contributing and non-contributing safe bases for belief:

> B is a safe basis upon which to believe that P just in case in all the most similar worlds in which B is true, P is true too. B is a contributing safe basis upon which to believe that P just in case, in addition, in all the most similar worlds in which P is false, B is false too. B is a non-contributing safe basis upon which to believe that P just in case, in some of the most similar worlds in which P is false, B is still true.[78]

To lean on the zebra case once more, when I believe that (P) those animals are zebras on the basis that (B) the animals look like zebras, B provides a contributing safe basis for belief that P. The nearest not-P worlds are not-B worlds, since presumably the nearest worlds in which it is false that there are zebras in the enclosure are worlds in which there is nothing that looks like a zebra. Furthermore, B also provides a safe basis on which to believe that (Q) those animals are not cleverly disguised mules, since all the nearby worlds in which B is true, Q is true also. However, this basis is of the non-contributing variety. The nearest worlds in which Q is false are worlds in which B is, in fact, true. That is, the nearest worlds in which the animals *are* cleverly disguised mules are worlds in which the animals do look like zebras.

When a basis for belief fails to contribute to the safety of the belief this is because the belief is safe purely in virtue of its content. What it means for a belief to be safe purely in virtue of its content is for the belief to enjoy a default level of safety in virtue of its modal profile. This is key to understanding Smith's account of transmission failure. Cases of transmission failure are those in which the inference fails to enhance the level of safety of the inferred belief because this belief already enjoys a supreme degree of safety in virtue of its content. That the animals are not cleverly disguised mules already enjoys a default level of safety in virtue of its content alone because the nearest worlds in which it is false are already very far from the actual

[78] Smith (2009, 170).

world. Likewise, that reality is mind-independent already enjoys a default level of safety in virtue of its content alone. This is not to say that one could never enhance the safety of these beliefs, merely that the inferences in question fail to do so. If one performed a DNA test on the animals, this would serve to enhance the safety of the belief that they are not cleverly disguised mules, thereby providing a contributing basis on which to believe, and thus a transmissive inference.

While Smith's account of contributing vs. non-contributing bases for belief provides an interesting way to think about failure of transmission of knowledge from within an externalist framework, it is again going to face similar problems to those raised for the reliabilist account of justification above. Once again, it seems to get the wrong result in *Twins*. To appreciate why, first assume that there is a safe basis on which to believe that (P) *this person is Sarah*. This includes the visual evidence of one who looks like Sarah, but in order for it to count as safe it needs to already include the proposition *this person is not Jessica*. Unless one can rule out that the person is Sarah's identical twin Jessica, the basis for belief that P will be unsafe as it will include nearby worlds in which the person is Jessica. So the basis for belief that P is something like (E) *this person looks like Sarah, and it is not Jessica*. This provides a safe basis on which to believe that P. However, if we accept Smith's description of transmissive inferences as those that provide a *contributing* safe basis for belief, we get the counterintuitive result that this inference is transmissive. The most similar worlds in which it is false that Q will presumably include worlds in which Jessica's business trip or holiday has been cancelled or cut short, and in which, crucially, E is false too (since the falsity of Q entails the falsity of E). As things stand then, the account of transmission failure developed by Smith gets the wrong result in this case. Given that we identified this case as one of the crucial cases for any theory of transmission failure to get right, this looks like a particularly bad result.

As before, the account also seems to run into trouble with beliefs in necessary truths. If my belief that P cannot fail to be false, then it seems as though the belief ought to be counted as safe in virtue of its content alone. As Smith says, "Certain beliefs—such as the belief that the animals before me are not disguised mules—enjoy a certain default level of safety in virtue of their content alone" (2009, 173). If the absence of normal worlds in which my belief that X is not a cleverly disguised mule means my belief is default justified independently of its basis, then it seems very much as though my belief that my belief that *the square root of 289 is less than 15* ought to be default justified independently of whatever I base that belief on, which is a counterintuitive result.

The same again is true of the last two problems raised above for the reliabilist account of transmission failure. There seems to be an absence of an independently motivated argument for why my belief in a Q-type proposition has to be based on an E-type proposition rather than a P-type. If the latter, then this seems to give a contributing basis for belief in the Q-type proposition after all. And lastly, the account once again seems insensitive to the possibility of unsafe bases for belief in Q-type propositions given that they are allegedly safe by default in virtue of their content alone. The upshot is that the safety-based account is on the same footing as the reliability-based account, facing four problems which jointly present a serious challenge to the attempt to diagnose transmission failure along the lines that Smith proposes.

It remains to be shown that the safety theorist does have the tools for meeting the minimal desideratum of capturing our intuitions in the most clear-cut cases of transmission failure such as *Twins*. By now the format should be fairly obvious. Cases such as *Twins* exhibit transmission failure because, in a very clear sense, they exhibit what I above labelled *positive dependence*. When I introduced this term earlier on for what Pryor calls type-5 dependence, this was in the context of a discussion of failure of transmission of *justification*. We are now thinking, not about justification, but about knowledge and safety. So, we need to understand how a safety-based account can accommodate this idea of positive dependence. The thought is going to be that an inference exhibits positive dependence if having a safe basis for belief in the conclusion is among the conditions that give you safe basis for belief in the premise. In other words, inferences that exhibit this type of dependence are such that the safe basis for the premise already includes a basis to believe the conclusion. It is clear that in *Twins* the basis for belief in the premise will exhibit this structure. There is a safe basis for belief that (P) *this person is Sarah* only if there is already a safe basis for belief that (Q) *this person is not Jessica*, given that Sarah and Jessica are identical twins that we have great difficulty in telling apart based on appearances alone. To see why this is, reflect that if the basis for belief that P does not already include a basis for belief that Q, the basis for belief that P will be unsafe. In such a case, the basis for belief that P would be the mere visual evidence (E) *this person looks like Sarah* and will allow for nearby worlds in which the person is in fact Jessica. In order for E to provide a safe basis on which to believe that P, it needs to be combined with a safe basis to believe that Q. This is how a safety-based account can accommodate the thought that the inference from P to Q exhibits positive dependence.

It is fairly trivial that an inference that exhibits positive dependence is non-transmissive. If my safe basis for belief that P already requires that I have a safe basis for belief that Q, then

I cannot acquire a safe basis for belief that Q via inference from P. There is a very obvious sense in which such an inference is epistemically circular. The account of transmission failure I am here offering is therefore not a particularly controversial one. It does at least have the merit that it captures our intuitions in the most clear-cut cases. What then does it say about the more controversial cases such as *Moore*? As with the simple reliabilist account of transmission failure I ended the last section on, this safety account of transmission failure predicts that the inference in *Moore* is transmissive. To see why notice that (P) *here is a hand* can be safely believed on the basis of (E) *it appears as though I have two hands*. All the nearby worlds in which E is true will be worlds in which P is true. Although there will be worlds in which E is true and P is false, these worlds can rightly be ignored for evaluations of safety. Because E alone is a sufficient basis on which to safely believe that P, the inference from P to (Q) *there is an external world* is transmissive. One could acquire, for the first time, a safe basis on which to believe that Q.

In terms of safety then, the Moorean inference is transmissive. This is not to rule out the possibility that the inference might fail to be transmissive of some other necessary condition for knowledge, simply that if the inference fails to transmit knowledge this will not be because it fails to transmit safety. One familiar point that the safety theorist can make regarding the Moorean inference is simply the same point made at the end of the above section on reliability. The safety theorist can appeal to the idea that the Moorean inference exhibits negative epistemic dependence and thus by entertaining sceptical doubts about the existence of the external world, one will rationally commit oneself to doubting both that one's visual evidence (E) supports belief in quotidian propositions such as P, and to doubting P itself.

3.7. Concluding Remarks

The account offered here is an attempt to understand transmission failure from within an externalist framework. In light of some suggestions that externalism is simply unable to accommodate intuitions of transmission failure I have shown that, in at least the most clear-cut cases, this is not the case. The resulting view is one according to which transmission fails in the clear-cut cases such as *Twins*, and *Coins*. The view predicts that there is successful transmission in *Moore*, and thus amounts to a kind of externalist, liberal, neo-Moorean response to scepticism. I have not developed this anti-sceptical view in any great detail beyond the basic structure outlined above, as it was sufficient to show that the very minimal externalist tools available can accommodate intuitions of transmission failure in the relevant cases. In the

following chapter, the anti-sceptical credentials of the view will be put to the test. I will consider two arguments either of which if successful would show that this view is unstable or incoherent. In the process of defending the account against these arguments, details of the view will be made clearer.

4.

4. Two Challenges to Externalist Neo-Mooreanism

The liberal externalist view advanced in the previous chapter amounts to a kind of neo-Moorean response to scepticism. The Moorean inference is transmissive so long as we construe warrant along externalist-friendly lines such as reliability or safety. In the present chapter, the view comes up against two major challenges. The first is specific to externalist responses to scepticism, while the second is specific to Moorean views in general. Either would suffice to undermine the view under consideration. The first challenge, due to Crispin Wright, is that externalist anti-sceptical views are inherently unstable—i.e., incoherent—because, when we pursue the implications of externalism sufficiently far, we find that we end up being forced to grant to the sceptic the very premise we set out to reject. The second challenge, due to Duncan Pritchard, is that all Moorean views rest on a mistaken assumption that anything that can be rationally inferred can be rationally believed.[79] The view is defended against both challenges, thereby strengthening the case that the externalist does have a coherent, plausible story to tell as to how warrant for the premise of Moore's argument can transmit via inference to its conclusion as a response to external world scepticism.

4.1. Scepticism and Closure

Earlier on in chapter I. we briefly noted that the Moorean inference can be turned on its head at the service of a particularly threatening form of scepticism.[80] The closure-based sceptical argument contraposes the logic of the Moorean inference. Granting Moore his major premise, the sceptic reasons that, indeed, *if one knows some empirical proposition P, then one knows the denial of some relevant sceptical hypothesis Q*. But, continues the sceptical train of thought, since one cannot know that Q, one cannot know that P. Where Moore takes it as a common-sense fact that one knows that one has hands, the sceptic argues we have better reasons for thinking that we cannot know the denials of sceptical hypotheses. One cannot know, for example, that there is a mind-independent reality, or that one is not a brain-in-a-vat, or that one is dreaming, and so on. If the sceptical argument can establish that we have better reasons for thinking that we don't know the denials of sceptical hypotheses than Moore gives for thinking that we know ordinary empirical propositions, then the neo-Moorean response to scepticism fails.

[79] This is not Pritchard's way of stating his claim, but I take it to be in the general spirit of his non-belief hinge epistemology, the precise details of which will be spelled out below.
[80] Most of the material from section 4.1. to section 4.4. is taken from 'On the Alleged Instability of Externalist Anti-Skepticism' (Jope, M., forthcoming[b]).

Here is how one version of the sceptical argument goes. The argument begins by constructing a sceptical scenario that is subjectively indistinguishable from what we take to be normality but that would have the effect that almost all of our beliefs would be false. A standard example is the brain-in-a-vat scenario: I believe in accordance with how things seem to me that I am sat at my desk, but might I not in fact be a disembodied brain floating in a vat of nutrients being electrochemically stimulated into having the illusory experience of being sat at my desk? The initial sceptical thought is that I have no way to know that this scenario does not obtain given that, if it did, things would appear to me just as they do now. The next step is to notice that most of my ordinary, everyday knowledge about the world deductively entails the denial of the brain-in-a-vat hypothesis. Take for example the proposition that I have hands. If I have hands then it follows deductively that I am not handless, and in particular not a handless brain in a vat. So now we have an ordinary proposition (*I have hands*) that entails the negation of a sceptical hypothesis (*I am not a brain-in-a-vat*). Knowledge closure tells us that if I know I have hands, then I know I'm not a brain in a vat, and by contraposition if I *don't* know that I'm not a brain in a vat, then I don't know that I have hands. The argument thus runs as follows:

(1) I don't know that I'm not a brain in a vat
(2) If I know I have hands, then I know that I'm not a brain-in-a-vat
∴ I don't know that I have hands

And since practically all of our everyday knowledge entails the denial of some sceptical scenario or other, this formula can be repeated for everything we take ourselves to know. Letting P stand for an everyday ordinary proposition, SH stand for a sceptical hypothesis the denial of which is entailed by P and K stand for knowledge, the general schema of the closure-based sceptical argument is as follows

The Closure-Based Sceptical Argument[81]

(1) ~K~SH	(I don't know that SH is false)
(2) K(P)→K(~SH)	(If I know that P then I know that SH is false)
∴ ~K(P)	(I don't know that P)

[81] As we'll see in the following chapter, it is possible to restate the problem in the form of a paradox rather than an argument, which some authors take to be a more revealing way to think about the problem.

The argument is a powerful one because, while it isn't easy to point out where it goes wrong, its conclusion is so clearly unacceptable. It is hard to overstate the implications here. Sceptical hypotheses that target a variety of epistemic domains—knowledge of the past, knowledge of the future, knowledge of other minds, etc.—can be easily conjured up and in conjunction with closure entail that knowledge of ordinary propositions in those domains is impossible. For example, if I cannot know there exist other minds besides my own, then given closure I cannot know any proposition that entails this such as that my friend is in pain or that the president is upset. If I cannot know that the world did not just spring into being a moment ago complete with all my memories and apparent traces of a deeper past, then I cannot know that I went hiking last weekend. There seems to be no in principle barrier to constructing sceptical hypotheses like this for each and every domain of inquiry, meaning the structure of the above argument can be applied to each of these domains to demonstrate knowledge is impossible in each.

Just as with the Moorean argument, one possible response here is to grant the sceptical premise (1) that we cannot know the denials of radical sceptical hypotheses but limit the sceptical import by rejecting closure. Doing this would remove the support for premise (2) of the argument, thereby blocking the move to the argument's conclusion. But as we saw in the first chapter in relation to the Moorean argument, rejecting closure seems like it generates more problems than it solves. Another option is to reject premise (1) by showing that the sceptic is in no position to affirm that we cannot know—or otherwise have warrant to believe—the denials of radical sceptical hypotheses. Anti-sceptical responses of this variety need to show how it is that we can, contra the sceptic, have warrant to believe that some sceptical hypothesis such as the brain-in-a-vat scenario is false. In the following sections of this chapter, I am going to explore one intuitively quite promising strategy for doing this. This response begins by pointing out that the sceptical case for premise (1) presupposes a particular internalist account of warrant or justification which many epistemologists would object to on independent grounds and which, once rejected, removes the support for (1), thereby blocking the sceptical argument. This strategy has come under attack from Wright who argues that it is self-undermining or incoherent. This incoherence leads Wright to conclude that externalist accounts of warrant are in no position to respond to the problem and is therefore key to understanding the motivation for pursuing a more internalist strategy for dealing with the problem. Wright's own preferred strategy is to reject (1) by developing an account of epistemic entitlement, according to which we are warranted by default in trusting in the denials of sceptical hypotheses. This strategy will

be assessed later on in chapter four, but for now we are going to focus on Wright's argument for the instability of the externalist anti-sceptical strategy.

4.2. The Anti-Sceptical Challenge

The sceptic needs to motivate premise (1) of the closure-based sceptical argument. The standard approach is to argue that where it is impossible to subjectively distinguish between competing hypotheses, we are unwarranted in believing either hypothesis. The common-sense hypothesis is that it appears to me that I am sat at my desk typing out these words because indeed I am, and my perceptual faculties are a reliable indication of this fact. A possible alternative hypothesis is the aforementioned brain-in-a-vat scenario in which it appears to me that I am sat at my desk because I am a bodiless brain in a laboratory, subjected to experiments in artificially generated perceptual experiences, against my knowledge. As bizarre and fantastical as the sceptical hypothesis may be, it seems logically compatible with my current subjective experience—indeed, when spelled out in the right way it logically entails that I would be having the subjective experiences I am currently having. The sceptic claims that since we are unable to discriminate between such 'good case' and 'bad case' scenarios, there is no warrant to believe that we are in the good case. An obvious reply to the sceptic is to challenge the assumption that we are warranted in believing that scenario A obtains over scenario B only if we can subjectively discriminate between A and B. This assumption smacks of internalism and thus the avoidance of radical scepticism by rejecting it presents one of the principal advantages of embracing externalist epistemology. This approach has been challenged by Wright who argues that if we pursue the implications of a thoroughgoing epistemic externalism, we find that the resulting view is "unstable" or "incoherent" (2008, 513-514). This point "has not generally been grasped with any clarity" we are told (513). And thus, if sound, Wright's argument would present a novel and potentially decisive argument against externalism, distinct from more familiar complaints that externalist anti-scepticism is merely philosophically unsatisfying.[82]

Recall again the sceptical premise (1), which says that, for some sceptical hypothesis SH, we cannot know that ~SH. If the motivation for (1) is simply the thought that ~*SH* and *SH* are subjectively indiscriminable, then progress can be made by rejecting this subjective indiscriminability requirement on warrant in favour of an externalist alternative on which (1)

[82] For a discussion of the standard criticism of externalist responses to scepticism see Bergmann (2006, 213-218).

comes out as false or unwarranted. Such is the challenge to externalism. Ultimately, Wright thinks the challenge cannot be met. He briefly sketches how the externalist might try to reject (1), rehearsing safety and reliability readings of *W* on which (1) turns out to be unwarranted, but despite some initial plausibility he thinks the resulting view is unstable (2008, 513-514). I in the following sections of this chapter I will show that Wright's argument fails to deliver this result because in rehearsing the externalist view it slips in certain internalist assumptions about evidence and it is precisely these hidden assumptions that generate the instability, rather than anything inherent to externalism itself.

4.3. The Case for Externalist Anti-Scepticism

Consider how an externalist might try to meet the challenge by rejecting the subjective indistinguishability requirement on warrant in favour of a more externalist-friendly conception of when there can be warrant to believe a proposition. One strategy would be to appeal to an epistemic modal condition such as safety, which says that a belief is safe so long as it is true in the actual world and all or most nearby possible worlds too.[83] By definition, the denials of radical sceptical scenarios are highly modally robust and thus turn out to be excellent candidates for safe beliefs, provided they are true. This means that on a safety construal of warrant, if we are in the good case (~*SH*) then there can be warrant to believe that we are, i.e. *if ~SH then W~SH.* Crucially, whether there can be warrant to believe ~*SH* now has nothing to do with our capacity to subjectively discriminate between ~*SH* and *SH.* Whether there can be warrant to believe that ~*SH* now has everything to do with whether ~*SH*.

Wright asks us to consider the "intuitive operator" *it is not to be ruled out that*… (511). Nothing else is said about this operator beyond its intuitiveness, but since this is a rehearsal of an externalist view, we are presumably meant to interpret it in externalist-friendly terms (more on this below). We may presume, Wright says, that it is not to be ruled out that ~*SH*. This claim ought to be unobjectionable—even the sceptic should allow that *it is an epistemic possibility* that we are in the good case. Moreover, given that *if ~SH then W~SH*, we are now in a position to conclude that *it is not to be ruled out that it can be safely believed that ~SH.* In other words:

(2) *it is not to be ruled out that W~SH*[84]

[83] This formulation of safety is overly simplistic and will be prone to well-known counterexamples. However, I stick to it for the reason that it is the formulation that Wright uses (511) and that the differences between it and various refined formulations ought not matter for present purposes.
[84] Wright uses the label (*) for what I am calling (2).

And this, Wright thinks, is exactly what the anti-sceptic needs in order to reject the sceptical premise (1). After all, one way we can read (2) is *it is not to be ruled out that* ~(1). Having concluded (2) we have shown that the sceptic is in no position to affirm (1). Thus, by rejecting the subjective discriminability requirement on *W* in favour of an externalist safety requirement we have easily shown that there is no way for the sceptic to motivate a key premise in their argument, namely (1). Hooray for externalism.

4.4. Wright's Instability Argument Against Externalism

While the above picture seems like a prima facie plausible route to meeting the challenge, Wright claims that it is ultimately unstable. To appreciate why, Wright asks us to reflect that in order to take ourselves to be in a position to affirm (2), we must also take ourselves to be in a position to affirm

(3)~*WSH*[85]

the claim that there is no warrant *for* the sceptical hypothesis (2008, 513-514). A commitment to (2) is a commitment to (3) because *WSH* entails ~*W*~*SH*. In other words, the externalist must deny that one who believes that *SH* does so safely because if we do safely believe that *SH* then we do not safely believe that ~*SH*, which would thus vindicate the original sceptical premise (1).

The problem that Wright wants us to appreciate is that (3) ought to be just as objectionable to the externalist as the original sceptical premise (1): "the externalist is committed to regarding [3], externally construed, as just as inappropriate, and for the very same reasons, as the sceptical premise that we cannot warrantedly suppose that *SH* does not obtain" (2008, 514). Wright thinks that the externalist has no better assurance that there cannot be warrant for *SH* than the sceptic did that there cannot be warrant for ~*SH*. This is because the safety-based rationale for rejecting (1) allegedly works equally well as a rationale for rejecting (3): the externalist should allow that "we may […] warrantedly suppose that *SH* does obtain—if, for example, the belief that it does is, alas, safe: is true in all nearby worlds" (2008, 514). The resulting picture thus looks to be unstable and leads Wright to conclude that "externalism

[85] Wright uses the label (**) for what I am calling (3).

actually provides *no* coherent motive for" (2) "and thus no coherent motive for repudiating the original sceptical premise (1)" (*ibid*.).

Wright does not explicitly rehearse the corresponding rationale for the rejection of (3), beyond these brief remarks, but it is easy to construct by adapting the rationale he gives for rejecting (1). Recall the starting point of the argument for rejecting (1), namely the appeal to the innocuous claim *it is not to be ruled out that ~SH*. The parallel argument for rejecting (3) will thus begin with the corresponding claim that *it is not to be ruled out that SH*. And again, if *SH* is true then it could not easily have been false and thus could be safely believed. In other words, *if SH then WSH*. Finally, putting this all together, we can conclude that

(4) *it is not to be ruled out that WSH*

which is as bad for (3) as (2) is for (1). Effectively, (4) can be read as saying *it cannot be ruled out that* ~(3). Though this is not Wright's way of making the point, the problem is that the externalist is, in a sense, committed to a kind of Moorean absurdity: *3 but it cannot be ruled out that ~3*.

The argument for (4) is perfectly analogous to the argument Wright rehearses on behalf of the externalist for (2). Each step is the same as before, the only difference being the two different starting points: while the first began by assuming that we cannot rule out that we are in the good case, the second begins by assuming that we cannot rule out that we are in the bad case. Aside from these different starting points, the moves in both arguments are identical. Thus, insofar as the externalist appeals to the safety-based argument for (2) they are (allegedly) obliged to grant the reasoning in the parallel argument for (4). And yet, (4) is as much of a sceptical premise as (1). In fact, in precisely the same way that (2) entails (3), so does (4) entail (1), which means the externalist argument for rejecting (1) ultimately leads round in a circle, forcing us to concede the very premise we set out to reject. Ergo, the whole externalist picture is incoherent. Or so Wright's reasoning would have us believe. In fact, there is a critical mistake in all of this. Contrary to Wright's presentation of externalism, there is in fact no rationale for (4) which the externalist need accept.

4.5. Where the Instability Argument Goes Wrong

Wright's introduction of the *not to be ruled out that* operator has some questions hovering over it, one of which is the issue of what it would take to rule something out. This issue could be

framed in externalist, internalist, or in neutral terms. Though it is worth repeating, if this operator is to be understood internalistically, then Wright's argument would turn out to fail. After all, the argument purports to show that when we "pursue the implications of a thoroughgoing externalist construal of the relevant epistemic operator sufficiently far" the resulting externalist view is unstable (515). Hence, in pursuing such implications it is crucial that all epistemic operators appealed to work in ways that the externalist will accept; crucial that at no point in the argument are we falling back on an internalist conception of warrant or an internalist conception of what it would take to rule something out.

A lot now hinges on what, precisely, is meant by this *not to be ruled out that* operator. Here's a plausible first pass: what allows us to rule something out is our *evidence*. If *P* is not to be ruled out, then *P* is compatible with a given body of evidence. Likewise, if *P is* to be ruled out, then *P* is not compatible with a given body of evidence. So, the *not to be ruled out that* operator is satisfied in cases where there is no evidence against the proposition in question. Another way to put this is in terms of epistemic possibility. Where Wright says *it is not to be ruled out that P*, we may plausibly take him to mean *it is epistemically possible that P*. And again, propositions that are not epistemically possible are those propositions that are ruled out by our evidence.

Evidence is thus a key notion that is implicit in the instability argument at the point at which the 'not to be ruled out that' operator is introduced. We have already noted that it is crucial to the soundness of the instability argument that this operator is being understood in an externalist-friendly manner. The pertinent question is then whether the notion of evidence in play is acceptable to the externalist or whether it depends on presupposing some internalist notion of evidence. In both internalist and externalist traditions, various competing conceptions of evidence are possible. One key difference between the two is that internalist accounts of evidence will predict that subjects in subjectively symmetrical pairs of cases will not differ in terms of their evidence, regardless of how else they differ. So, in 'good' and 'bad' sceptical scenarios subjects share the same evidence. Most externalists, on the other hand, will predict that subjects in 'good' and 'bad' cases do not have the same evidence.

In what follows it will be useful to have a particular externalist account of evidence in mind. The argument that follows will work on a variety of different externalist accounts that hold that evidence is factive, but it will be helpful to be more specific than this in order to try to run the instability argument against it. One account that has enjoyed increasing popularity in recent years equates evidence with knowledge. On such an account, *P* is part of one's

evidence if and only if one knows that *P*. Call this the *E=K thesis*.[86] In line with the discussion of the *not to be ruled out that* operator above, on this account, what it takes to rule out *P* is knowledge of something that entails ~*P*. It follows that on such an account, if one is in the ~*SH* scenario, then one's evidence set contains a plethora of ordinary, empirical propositions such as *there are chairs, there are tables,* etc. Conversely, in the *SH* scenario, one's evidence is woefully impoverished; one's evidence set does not contain propositions such as *there are chairs, there are tables.*

Recall the first steps of each of the arguments for rejecting, respectively, (1) and (3):

(a) it is not to be ruled out that ~SH
(b) it is not to be ruled out that SH

For the instability argument to go through, it is necessary that both arguments are acceptable to the externalist, and thus that each of these first premises, (a) and (b), are acceptable. It is not to be ruled out that the sceptical hypothesis is false, and it is not to be ruled out that the sceptical hypothesis is true. The intuition has to be that our evidence does not allow us to rule in either direction. This intuition is suspect. It depends on an internalist conception of evidence, according to which our evidence is the same across *good* and *bad* pairs of cases. While (a) is true on both internalist and externalist accounts of evidence, (b) is false on any factive conception of evidence because our evidence entails ~*SH* and thus it *is* to be ruled out that *SH*.

First to demonstrate that (a) is true. Recall that what would allow us to rule out that ~*SH* would be evidence for its negation, i.e. knowledge of something that entails *SH*. It is easily shown that ~*SH* is not to be ruled out by showing that in both *good* and *bad* cases, there is no evidence for (knowledge of) *SH*.

Argument for (a)

(i)	$SH \rightarrow \sim K(SH)$	(by stipulation of the sceptical case)
(ii)	$\sim SH \rightarrow \sim K(SH)$	(trivial by factivity of *knowledge*)
(iii)	$\Box(SH \vee \sim SH)$	(*law of excluded middle*)
∴	$\sim K(SH)$	(i, ii, iii)

[86] For a statement of *E=K* see Ch. 9. of *Knowledge and its Limits* (Williamson, 2000).

Either *SH* is true or ~*SH* is true (by law the excluded middle). If ~*SH* is true, then trivially one does know that *SH* (because knowledge is factive). If *SH* is true, then, by stipulation of the sceptical case, one does not know that *SH*.[87] Either way, one does not know that *SH*, has no evidence for *SH*, and since we are understanding Wright's *not to be ruled out that* operator in terms of evidence, it is therefore not to be ruled out that ~*SH.*

Now to demonstrate that (b) is false on the externalist account. If it is not to be ruled out that *SH*, then there is no evidence for ~*SH*. While in the *bad case* this is trivially true, in the *good case* this is false.

Argument for (b)

(i*) ~*SH* → ~*K*~*SH*	(??)
(ii*) *SH* → ~*K*~*SH*	(trivial by factivity of *knowledge*)
(iii*) □(*SH* v ~*SH)*	(*law of excluded middle*)
∴ ~*K*~*SH*	(i*, ii*, iii*)

This argument is structurally identical to (a) and is therefore valid, however it fails to support its conclusion because premise (i*) is false. The externalist does not need to grant that in the good case one cannot have evidence that one is in the good case. In the good case (~*SH*) one enjoys a plethora of perceptual evidence that supports ~*SH*. That evidence is just one's total body of evidence—one's knowledge—virtually all of which entails ~*SH*. For example, my knowledge of the existence of objects in my immediate visual field such as chairs and tables entails that ~*SH*. Because this argument fails, the first premise of the safety argument for the rejection of (3) is false. Thus, the argument is unsound. The upshot of this is that the instability argument fails to demonstrate that pursuing externalism to its logical conclusion leads to incoherence. That argument depended on there being parity between a safety argument for (2) and a safety argument for (4). There is no such parity because, while the starting point of the former argument (a) is uncontroversial, the equivalent starting point of the latter argument—(b)—is false on an externalist conception of evidence.[88]

[87] The orthodox view in epistemology assumes that if one is in the sceptical case, then one cannot know this. If I am a brain-in-a-vat, then I cannot know that I am, even were I to believe it. For one rather radical and recent challenge to this orthodoxy, see Magidor (2018). For a reply to Magidor see McGlynn (2018).

[88] While the focus here has been on a knowledge-first conception of evidence, this point will extend to other externalist accounts of evidence as well. Importantly, any account according to which evidence is factive could produce the same result.

Would there be parity between the two arguments on an internalist account of evidence? This is an open question and will differ depending on the particular internalist account one favours. All internalist accounts are committed to the thought that, whatever evidence subjects have for their beliefs, it is the same across *good* and *bad* sceptical cases. Whether this evidence allows subjects to rule out *SH* or not depends on what a specific internalist account says about our evidence. However, one thing is clear: in motivating the instability argument, Wright takes it to be intuitively plausible that we cannot rule out *SH*, and this intuition does depend on an internalist construal of evidence. A construal, according to which, our evidence does not allow us to rule out *SH* since we would have the same evidence even if *SH* were true. If that is the intuition behind the supposedly plausible (b) that I have identified as necessary in generating the instability argument, then it is clear that the instability that Wright claims to highlight in the externalist position comes not as a result of anything specific to externalism, but rather from an underlying internalist assumption about evidence.

4.6. The Second Challenge to Externalist Neo-Mooreanism: Non-Belief Hinge Epistemology

The Neo-Moorean view advanced in chapter III. amounts to the claim that by appealing to an externalist conception of warrant—either in terms of reliability or safety—we are able to vindicate Moorean anti-sceptical reasoning.[89] This vindication of Moorean reasoning will fail if Wright's argument for the instability of externalist responses to scepticism is sound. In response I have shown that Wright's argument rests on some mistaken conceptions about externalist conceptions of evidence and that once these mistakes are addressed, the alleged instability dissipates. With the externalist conception of warrant on the table, it looks like we can, contra the sceptic and in line with Moore, have warrant for the denials of radical sceptical hypotheses. So far this vindicates the possibility of a safe or reliable basis for belief in the conclusion of Moore's inference. So far so good. But the view now faces another substantial challenge. The conclusion of Moore's inference is not any ordinary kind of proposition. It is what we have varyingly been referring to as a 'cornerstone', or 'authenticity condition', or simply a 'hinge' proposition. Earlier on in chapter II. we saw that hinge epistemology holds that it does not make sense to rationally doubt hinges. But suppose that we thought that for very similar reasons, nor does it make sense to *believe* them. If this claim can be substantiated,

[89] Most of the material from section 4.6. to 4.9. is taken from 'Closure, Credence and Rationality: A Problem for Non-Belief Hinge Epistemology' (Jope, 2019).

then arguments to the effect that we can have a robust modal or reliable basis for believing hinges are moot.

As noted briefly in chapter II., the metaphor of a 'hinge' proposition or commitment is due to Wittgenstein. He believed that certain kinds of propositions, about which one is optimally certain, play a special kind of role that sets them apart from the ordinary channels of inquiry. Just as a door cannot turn without its hinges staying put, so inquiry cannot proceed without certain commitments standing fixed. A variety of competing accounts of hinge epistemology exist in the literature, distinguished from each other by their responses to several issues such as whether hinges are propositional, knowable, dubitable, empirical, fact-stating, justifiable and rational.[90] Duncan Pritchard labels his own version of hinge epistemology a *non-belief* account of hinges, a key feature of which is the surprising claim that hinges are commitments that we cannot believe despite our being certain of them (2016a, 90-94; 2016b, 78-82).

According to the 'non-belief' reading of hinges, these are propositions that lay beyond the realm of *rational grounding*, so just as it makes no sense for the sceptic to offer grounds for doubting hinges, it does not make sense for the anti-sceptic to offer grounds for believing them. Such propositions have a special kind of status in our epistemic lives which rationally precludes the possibility of our believing them. If the non-belief reading of hinges is correct, then hinges can neither be known nor justifiably believed and the Neo-Moorean view under consideration is guilty of committing a kind of category mistake: the concept of rational belief is simply inapplicable to hinges.

In the remainder of this chapter, I examine Pritchard's argument for this claim, to shed light on a critical assumption at work in the argument and to demonstrate that the assumption is flawed. In section 4.7. I outline Pritchard's response to closure-based scepticism and the non-belief reading of hinges which underpins it. In section 4.8. I show that Pritchard's argument for the non-belief reading relies upon the assumption of a principle about the nature of rational support which, depending on how it is unpacked, either leads to irrational doxastic attitudes or is inapplicable to closure-style inferences. In section 4.9. I consider whether a modified version of Pritchard's account might avoid this problem but conclude that ultimately it cannot.

[90] For a recent taxonomy of different hinge epistemologies and their positions on each of these issues see Coliva (2016).

4.7. Pritchensteinian Rational Grounding

Rather than framing the closure-based sceptical problem as an argument with a sceptical conclusion—as it was presented above in section 4.1. Pritchard frames the closure-based sceptical problem instead as a paradox. To begin with he offers the following specific formulation of the closure principle for rationally grounded knowledge:

> *The Closure Principle*
> If S has rationally grounded knowledge that P, and S competently deduces from P that Q, thereby forming a belief that Q on this basis while retaining her rationally grounded knowledge that P, then S has rationally grounded knowledge that Q (2016b, 71).

Thus formulated, the principle looks eminently plausible. The paradox arises by combining the closure principle with two further claims:

> *The Closure-Based Sceptical Paradox*
> (I) One cannot have rationally grounded knowledge of the denials of radical sceptical hypotheses
> (II) The closure principle
> (III) One has a large body of rationally grounded knowledge of the external world

In presenting the problem of closure-based scepticism as a paradox rather than an argument, Pritchard is able to demonstrate that the problem is not a dialectical standoff with an imagined sceptical opponent, but rather "a deep tension within our own folk epistemological concepts" (2016a, 15). We feel inclined to grant that knowledge of the denials of radical sceptical hypotheses is not possible. But we also feel inclined to grant that ordinary, everyday knowledge is possible. And yet, the closure principle seems to entail that these two claims cannot simultaneously be true. Traditional responses to the paradox thus come in three varieties, equivalent to the rejection of one of each of the three claims in the paradox. Rejecting (III) is tantamount to embracing radical scepticism. Rejecting (II) means giving up an extremely plausible epistemological principle.[91] And rejecting (I) would mean embracing a form of Mooreanism, which many epistemologists—Pritchard included—find to be intellectually unsatisfying. Pritchard points out that, to the extent that we find each of these three claims

[91] A majority of epistemologists find this option unsatisfying.

highly intuitive, any strategy that seeks to respond to the paradox by overturning one of them would be revisionist (2016b, 72).

Pritchard argues there is a fourth way out of the paradox that avoids such revisionism and that, for this reason, is to be preferred from its alternatives. Instead of revising our epistemological theory by overturning one of the three horns of the trilemma, Pritchard argues that, despite appearances, these claims are not in fact inconsistent. This is because, as we will see, Pritchard thinks that the closure principle is inapplicable to the denials of radical sceptical hypotheses, and thus it is a mistake to think that (III) and (II) jointly entail the denial of (I)—or likewise that (I) and (II) entail the denial of (III). In order to show how this surprising claim could be true, he appeals to some remarks from Wittgenstein's notes on scepticism.[92] The key thought behind this approach is that our anti-sceptical commitments are not in the market for rational evaluation in the way that both sceptic and Moorean strategies assume. The sceptical strategy assumes that our anti-sceptical commitments are open to rational doubt, while the Moorean response assumes the possibility of a rational basis for belief in our anti-sceptical commitments.[93] Pritchard insists, however, that our anti-sceptical commitments belong to a special class that, following Wittgenstein, he calls "hinge commitments" (2016a, 66-73; 2016b, 79). Hinge commitments are arational standing certainties that are not the object of investigation—whether *a priori* or *a posteriori*—but are rather part of the presupposed background, or "framework", relative to which rational evaluation takes place (2016b, 78). Pritchard claims that because our hinge commitments are neither the result of a rational process nor subject to rational considerations, both the sceptical and anti-sceptical strategies for dealing with the paradox are misguided (2016a, 65-66; 2016b, 78). In sum, we can endorse the closure principle unreservedly and yet not be forced to reject either (I) or (III) once we come to appreciate the kinds of limits to rational evaluation that Wittgenstein uncovers in his remarks on scepticism.

The proposal that Pritchard puts forward holds that hinge commitments are, at one and the same time, neither in the market for rational doubt nor for rational belief (2016a, 63-66; 2016b, 78-84). In what follows I will be focusing in on the latter of these two claims regarding

[92] The notes referred to here are those posthumously published as *On Certainty* (1969).

[93] Pritchard includes among 'Moorean' anti-sceptical strategies the kind of 'entitlement' view defended by Crispin Wright (2004, 2014). This is a somewhat unusual way to categorise the views on offer given that Wright explicitly positions his view against Mooreanism. Furthermore, Wright is clear that he does not think there can be a rational basis to *believe* in anti-sceptical hinge propositions such as *there is an external world* (according to Wright, we are warranted to *trust*, rather than believe, in such). However, the point that Pritchard seems to want to make here is that, like the standard Moorean anti-sceptical strategy, Wright's entitlement strategy does take there to be a rational basis for our anti-sceptical hinge commitments. In contrast, the view that Pritchard endorses denies that there can be a rational basis for our hinge commitments.

the impossibility of a rational basis for belief. Hinge commitments are commitments to those propositions about which we are *optimally certain* (Pritchard, 2016b, 78). This includes propositions such as 'I have two hands' as well as 'there is an external world'. Given their optimal certainty, hinge commitments "cannot be subject to rational doubt" (*ibid*.). This is because doubt about that which is optimally certain would throw into question "one's entire system of beliefs, and thus the very putative rational basis of the doubt itself" (2016b, 79). Ordinarily, one might think that this means these propositions enjoy a uniquely solid epistemic basis. Not so, says Pritchard: "it does not follow that these hinge commitments have a special rational grounding, but rather just as they cannot be rationally doubted, so they cannot be coherently thought of as rationally grounded either" (*ibid*.). Propositions about which one is optimally certain cannot, according to Pritchard, have a rational basis because "to conceive of [a] proposition as rationally grounded is to suppose that the rational grounds are *more certain* than the proposition itself" (2016a, 65; emphasis added). Since nothing can be more certain than that which is optimally certain, our optimally certain hinge commitments cannot be rationally grounded. This is to say that "one cannot make sense of a rational basis for *belief*" in hinges (2016b, 79; emphasis added). For this reason, Pritchard calls his view the non-belief reading of hinge commitments (2016a, 91).[94] And since hinges cannot be believed, neither can they be known. Thus, while the indubitability aspect of the view separates it from sceptical responses to the paradox, the non-belief aspect separates it from 'Moorean' responses.

One might think that the upshot of the above is that Pritchard is forced to deny closure. After all, if our hinge commitments aren't in the market for rationally grounded knowledge, then it looks like it ought to be very easy to generate counterexamples to closure. In fact, Pritchard maintains that his view can embrace the closure principle unreservedly. He argues that what the view really entails is that "the closure principle is simply inapplicable to our hinge commitments" (2016b, 81). The reason being that the closure principle applies only to cases where a belief is formed via deductive inference, whereas according to the non-belief reading this is "simply impossible" (2016a, 86). The closure principle thus cannot be used to motivate scepticism, nor can it be used to defend Mooreanism.

4.8. Against Pritchensteinian rational grounding

[94] In contrast to what he calls the externalist, entitlement, and non-propositional readings (2016a, 73-88).

Notice that Pritchard's argument depends upon the assumption that a proposition is rationally grounded only if the rational grounds are "more certain" than the proposition itself (2016a, 65). The claim that hinge commitments cannot have a rational grounding—and thus cannot be thought of as *beliefs*—fully depends upon this assumption. Without it, it is unclear why the mere fact of the optimal certainty of hinges would preclude them from counting as rationally grounded; more traditional approaches would hold that a proposition about which one is optimally certain ought to possess an extremely solid rational basis. The argument for the arationality of hinges thus depends on assumption of the following principle:

> *The Pritchensteinian Rational Grounds Principle*
> A proposition p is rationally grounded only if the rational grounds are more certain than the proposition itself[95]

For this principle to do the work that Pritchard needs it to do, it ought to apply specifically to cases where the rational grounds are themselves other propositions. Given that we are interested specifically in closure, and thus in whether a *proposition* can provide rational grounds on which to believe another proposition (as opposed to, say, an experience, i.e. a perceptual experience, providing grounds to believe a proposition), it is necessary that the rational ground(s) in question is a proposition(s). We can thus modify the Pritchensteinian Rational Grounds Principle in the following way:

> (PG1) One proposition P can count as a reason for believing another proposition Q only if P is more certain than Q

This assumption appears crucial to Pritchard's argument that the closure principle is inapplicable to our anti-sceptical hinge commitments. Without it, we have no reason to suppose that one cannot acquire a rationally grounded belief in a hinge by performing a competent deduction from a rationally grounded belief. There are some very serious problems that this principle faces which we shall see shortly, but first I want to note some substantive commitments implicit to PG1.

---

[95] The principle stated here is a paraphrase of the following quote: "to conceive of [a] proposition as rationally grounded is to suppose that the rational grounds are *more certain* than the proposition itself" (2016a, 65; emphasis added). Pritchard again appeals to the same assumption further down on the same page when he says that "If one doubts that one has two hands, then one ought not to believe what one's eyesight tells one, since this is *no more certain* than that one has two hands, which is in doubt" (*ibid*.; emphasis added).

PG1 presupposes a gradable notion of 'certainty'. There is a strict sense in which 'certainty' is sometimes used as an all-or-nothing notion, meaning something like credence 1. But there is another sense in which 'certainty' is used in a gradable manner, as when we say things like 'fairly certain' or 'absolutely certain'. Clearly, Pritchard is using 'certainty' in the latter sense when he speaks of propositions being 'more', 'less' or 'optimally' certain (2016a, 64-65). Thus PG1, which we have extracted from Pritchard's argument, uses a gradable notion of certainty. In addition to this, PG1 also entails that a proposition P can be a reason for believing another proposition only if P is certain (P cannot be *more certain* than Q without being certain). This is a substantive claim, given that many epistemologists will grant that that belief does not require certainty[96] and that if it is rational to believe P then it is rational to believe (at least some of) P's consequences. PG1 is forced to deny this liberal conception of belief according to which it can be rational to believe things about which one is less than certain. On the other hand, some epistemologists maintain that belief entails certainty, and thus on such a view PG1 is perfectly acceptable.[97] I note this merely to flag PG1's substantive commitment to a strong conception of belief that entails certainty.

One final clarificatory point. Degrees of belief are commonly measured in terms of credences or degrees of confidence. Pritchard, on the other hand, seems to measure rational belief in terms of degrees of *certainty*.[98] In what follows, I will be assuming that degrees of confidence and degrees of certainty map onto each other in a fairly straightforward way. I will assume that both are a measure of credences such that to become more and more confident of a proposition is to become closer and closer to certainty. I will not argue for this assumption but will note the following point. The only other alternative is to radically disconnect the notions of degrees of confidence and degrees of certainty in such a way that what Pritchard means by a belief being grounded in a particular degree of certainty is not to be understood in terms of a high degree of confidence or particular credence. I do not think this alternative has much going for it, but I will flag it as a possible avenue for avoiding the dilemma below which assumes a close connection between degrees of certainty and degrees of confidence.

We are now in a position to explore the dilemma for Pritchard's account. As it stands, PG1 can be read in at least two different ways. On one of these readings, it is completely implausible, yet on the other reading it is simply inapplicable to closure-style inferences. To

[96] See, for example, Foley (1992).
[97] See, for example, Dodd (2016).
[98] Recall the Pritchensteinian Rational Grounds Principle and the quoted passages in support of Pritchard's endorsement of it.

see what I mean by this, notice that we may distinguish between two ways in which one proposition can provide a rational basis for another proposition, relating to the kind of support relation between the two. Say that one proposition can provide a *deductive* rational basis for another so long as the former deductively entails the latter. Likewise, say that one proposition can provide a *non-deductive* rational basis for another so long as the former non-deductively—i.e. inductively or abductively—supports the latter. PG1 can thus be interpreted in two ways, relating to these two types of support relation.

We can rule out the non-deductive reading of PG1 very quickly. Closure-style inferences by definition involve deductive inferences—e.g. *here is a hand, therefore the external world exists*.[99] Thus, on assumption that PG1 is a principle about strictly non-deductive inferences, it is inapplicable to the issue of whether (deductive) closure-style inferences can serve to provide a rational basis on which to believe our anti-sceptical hinge commitments.

This leaves only the option of a deductive reading of PG1. However, if we take a deductive reading of the rational support relation, PG1 is completely implausible. It clashes with a basic principle of rationality. It demands of us that we are more certain of the antecedent of a conditional than of its consequent. It should be clear from the discussion in chapter II. on problems for so-called 'weak' accounts of entitlement to trust why this demand leads to irrationality. Recall that we noted there that rationality places certain demands on how our credences across inferences ought to be spread. We first noted that the following is a consequence of the axioms of probability:

If P→Q then $Pr(P) \leq Pr(Q)$

which says that where P entails Q, the probability (*Pr*) of P cannot be higher than that of Q. This principle relates to objective probability, but it can easily be turned into a principle about rational credences if we assume that a rational agent's credences will satisfy the laws of probability. From this we were able to derive the following principle:

(RC) where P deductively entails Q, it is irrational to be more certain of P than of Q

[99] This infamous example goes back to G. E. Moore's "Proof of an External World", in which Moore is very clear that the relevant inferences need to be deductive (2000, 26-27).

Once again, exceptions to the principle are those in which the subject is ignorant of the entailment. Bracketing these, RC seems like a watertight principle of rational credences. The problem now of course is that PG1, when read with a deductive reading of the rational support relation, is incompatible with RC. If credences must not drop across deductive entailment, then it cannot be a requirement of rational support that they do so. So, to the extent that RC seems fundamental to our conception of rationality, the deductive reading of PG1 is a non-starter.

A possible objection to this line of reasoning is the following. RC is a principle that concerns our *rational* attitudes. However, the very notion of a hinge commitment is of a type of attitude that is, by definition, *arational*. If hinge commitments lay beyond the scope of the commitments of rationality in the way suggested by hinge epistemologists, then perhaps principles like RC get no purchase on them.

The problem with this objection is that it is begs the question. Hinge epistemologists argue that hinges are arational on the grounds that there can be no rational basis to believe a hinge. We want presently to investigate what basis there is for this claim and thus in doing so we cannot help ourselves to the assumption that hinges are arational, since that is precisely the question we are attempting to settle.

Another possible objection to the above argument is that it presupposes the truth of *probabilism*—the view that our credences ought to satisfy the axioms of probability. Probabilism is a contentious view in epistemology and for anyone who is unpersuaded by its truth, the forgoing may seem relatively unpersuasive.[100] I do not think that the argument ultimately rests on an assumption of probabilism but before getting to that, I first want to draw the reader's attention to the following. One reason for objecting to understanding our epistemic practices in terms of the probability axioms is that this mischaracterises the subject matter of epistemology as an inquiry into the nature of perfectly rational agents, whereas human beings are inherently imperfectly rational. This objection maintains that probabilism is in some sense too demanding—requiring doxastic attitudes that are fine grained, whereas our doxastic states are coarse-grained. One thing that I think is worth bearing in mind is that Pritchard and Wittgenstein seem to have a very specific notion of rationality in mind—one that I am not convinced this objection will apply to. This is to say that the Pritchensteinian account of rationality is explicitly not an account of the limits of rationality qua *human* rationality. The limits to rationality that Pritchard and Wittgenstein are trying to convince us of are "not because of some incidental lack on our part, but rather [reflect] the very nature of what is involved in

[100] For a critical discussion of the main arguments in defence of probabilism see Hajek (2008).

rational evaluation" (2016a, 4). Wittgenstein's point, Pritchard assures us, is a matter of the "logic" of rationality (*ibid*.). And I suspect that whether or not it makes sense to understand our actual epistemic practices in terms of something like the probability axioms, it perhaps still makes sense to understand an idealised form of rational evaluation this way.

A more robust response to this objection is that even if we disregard probabilism altogether, the essential problem remains that it is still inherently irrational to be more confident of the antecedent of a conditional than of its consequent. [101] Earlier on I argued for RC by appealing to the probability axioms together with the assumption that a rational agent's credences will satisfy those axioms. But suppose we reject this assumption and maintain that credences are essentially non-probabilistic. It will not follow from this that rationality now permits higher degrees of certainty in the antecedents of conditionals than their consequents. I am not sure how best to argue for this but consider the following line of thought. To be more certain of P than of Q is to believe that one possible state of affairs is P and ~Q. At the same time, to believe that P deductively entails Q is to believe that it is not possible that P and ~Q—this ought to follow given any standard definition of deductive entailment. Therefore, to be more certain of a proposition P than of something one believes to be a deductive consequence of P, is to believe a contradiction, namely: it is possible that [P^~Q] and it is not possible that [P^~Q]. On assumption that it is irrational to believe something one knows to be a contradiction, we have another way of showing why it is irrational to be more confident of the antecedent of a conditional of than of its consequent. The appeal to probabilism was one very precise way to illuminate this irrationality, but there are other ways to do this that do not depend on probabilism.

4.9. An alternative Pritchensteinian principle?

We have seen that the thesis that a proposition can be rationally grounded only if its rational grounds are more certain than it, leads to irrationality. This means that if there is indeed something about the optimal certainty of hinges that precludes their counting as rationally grounded there needs to be some other explanation for why this is.

Consider the following example that Wittgenstein and Pritchard discuss in relation to the claim that optimal certainty of a hinge precludes its counting as evidentially grounded:

[101] Again, to avoid confusion, this assumes that one grasps the conditional and the conditional is not false. There need be nothing irrational in being more confident of the antecedent of a conditional than its consequent where one has yet to grasp the entailment. Likewise, no irrationality need follow from being more confident in the antecedent of a false conditional.

> If a blind man were to ask me "Have you got two hands?" I should not make sure by looking. If I were to have any doubt of it, then I don't know why I should trust my eyes. For why shouldn't I test my *eyes* by looking to find out whether I see my two hands? *What* is to be tested by *what*? (OC, §125).

In discussing this case, Pritchard again assumes that something akin to PG1 is in play: "If one doubts that one has two hands, then one ought not to believe what one's eyesight tells one, since this is *no more certain* than that one has two hands, which is in doubt" (2016a, 65; emphasis added). But we've already seen why PG1 can't in general be an explanation for why hinge commitments are groundless. So perhaps there is a different way to capture what Wittgenstein is gesturing at with this example that doesn't assume PG1. After all, nothing in the example seems to *necessitate* Pritchard's point that what is going on is a failure of the basis to be *more certain* than the hinge proposition (though the passage permits this reading). An alternative explanation could be the following. Since hinges are already optimally certain, there is nothing that one could appeal to that would warrant an increase of certainty in one's hinges. If I am already maximally certain that I have two hands, of course it would be absurd to check by looking to see whether I do since there is no possible outcome of this investigation that would increase my certainty in my belief that I have two hands. We thus have a possible alternative to PG1:

> (PG2) One proposition P can count as a rational basis for believing another proposition Q only if P warrants an increase of confidence in Q[102]

To be clear, neither Wittgenstein nor Pritchard explicitly endorse PG2. The claim here is not an exegetical one. Rather, since PG1—which Pritchard does seem to endorse—fails to stand up to scrutiny, perhaps there is an alternative understanding of why it is the case that something's being optimally certain precludes it from counting as rationally grounded. PG2 looks like a potential alternative to PG1 that shows some promise of capturing why this might be. There are, however, good reasons to be suspicious of PG2 also.

Sometimes, our reasons for belief overdetermine the beliefs they support. And yet, they are reasons for belief nonetheless. It can easily be the case that I acquire a reason for believing

[102] I switch here to talking of degrees of confidence rather than degrees of certainty.

a proposition without there being an increase in the degree of confidence I am warranted in taking towards that proposition. This is the case when I already possess stronger reasons for belief prior to acquiring the additional reasons. Consider the following case. Suppose a legal setting in which it is being determined whether a defendant, Smith, committed a crime at a given place and time. An expert gives testimony to the effect that Smith's DNA matches that of the DNA found at the scene of the crime. Clearly, this is a reason for believing that the defendant was at the scene of the crime at some point. Even for Smith himself, who already knows that he was at the scene of the crime, the fact that his DNA has been found at the scene of the crime is still a reason for believing that he was there. And yet, this is not to say that it is a reason that ought to increase his confidence that he was there, since he already has a much stronger basis on which to believe that. This is in conflict with PG2, which entails that unless the discovery that his DNA was found at the scene of the crime increases Smith's warranted degree of confidence in his belief that he was at the scene of the crime, it is not a reason—a rational basis—for that belief. All the worse for PG2.

4.10. Concluding Remarks

To recap, we have seen that the externalist neo-Moorean view advanced in chapter III. faces at least two serious challenges. The first of these is that externalist responses to scepticism in general, if pursued to their logical conclusions are revealed to be unstable. If sound, this argument would represent a decisive case against externalist anti-sceptical strategies. However, it was shown that the argument for the instability of externalism rests on a mistaken assumption about how the externalist conceives of evidence. Crucially, if we allow for the possibility of factive evidence, then we can show how there can be warrant for the denials of radical sceptical scenarios without needing to concede to the sceptic the claim that this evidence leaves open the possibility that such hypotheses may, after all, be true. This argument therefore vindicates the possibility of an externalist neo-Mooreanism. Having established this much, we then turned to a separate challenge due to Duncan Pritchard which alleged to show that, no matter how we conceive of the warrantability of such bases for belief, the conclusion of Moore's inference belongs to a unique class of propositions—hinges—that lay beyond the scope of rational reasons or grounds. If Pritchard's argument is right then the optimal certainty of our anti-sceptical hinge commitments precludes a rational basis for belief in them, which would block any kind of neo-Moorean response to scepticism while giving us an entirely different story to tell about what is wrong with Moorean reasoning from the one given in chapter III. The non-

belief reading of hinges holds that it is a kind of category mistake to apply the notion of rational belief to propositions such as *there is an external world*, so there is no way to acquire a justified belief in the conclusion of Moore's inference. However, it was then argued that this view rests on a problematic conception of rational grounding—what I labelled *The Pritchensteinian Rational Grounds Principle*—that we have independent reason to want to reject. This principle effectively says that it is a condition on one proposition counting as a basis for belief in another proposition that the former is more certain than the latter. I have argued that insofar as this principle is intended to be a principle that applies to deductive entailments between propositions—and it better be if it is to apply to closure-style inferences—it is incompatible with some very basic principles about rationality. In light of this, I have tried to sketch an alternative understanding of why it could be that optimal certainty precludes the possibility of a rational basis for belief but found that this alternative proposal faces fatal problems of its own. Pritchard's argument therefore fails to support the claim that the denials of radical sceptical hypotheses cannot be rationally believed.

The externalist neo-Moorean view under consideration thus holds up against the two challenges considered here. These challenges promised to pose a more serious threat to the viability of this brand of response to scepticism than more familiar complaints that externalism in particular, and neo-Moorean views in general, provide a mere conditional response to scepticism, or are in some sense philosophically unsatisfying, or do not take the problem of scepticism seriously enough.[103] If successful either one of the two challenges addressed would suffice to demonstrate that externalist neo-Moorean views such as the proposal advanced in the previous chapter are quite literally *incoherent*. Happily, for fans of externalism we have seen that these challenges fail to stand up to scrutiny.

[103] For an extensive discussion of the complain that externalism provides a mere conditional response to scepticism see the discussion between Ernest Sosa (1994) and Barry Stroud (1994). For discussion of the complaint that externalist anti-scepticism is somehow philosophically unsatisfying see Bergmann (2012, §8.) and also Pritchard and Ranilli (2016).

5.

5. Theoretical Symmetry and Testimonial Conservatism

A great deal of the conservative-liberal debate focuses on the structure of *perceptual* justification.[104] Most of the examples focused on in the literature, and indeed in previous chapters here, are of this kind. The Moorean inference is of course one such case involving an inference whose premise is based on perception and whose conclusion is an authenticity condition for the method of forming beliefs based on perception. But it should also be clear that we can talk about the structure of justification of other epistemic sources or methods too, as well as authenticity conditions for reliance on those methods. One such example that we briefly touched on in chapter II. is an inference whose premise is based on memory, which depends in some way on the authenticity condition *memory is reliable*. Another interesting type of case to think about is an inference whose premise is based on testimony and which depends in some way on testimonial authenticity conditions. The focus of this final chapter will be on the relationship between conservative views about testimony and conservative views about perception. The reason for this focus is that, while most of the liberalism-conservatism debate has focused on the case of perception, there is an entirely separate body of literature on the epistemology of testimony, which in some clear sense mirrors a lot of what is going on in the former debate. The purpose of this chapter is to draw a link between these two literatures and to argue for the claim that a commitment to a certain view about the structure of testimonial justification entails a commitment to a corresponding view about perceptual justification. That is, what I will identify as *testimonial conservatism* entails a commitment to *perceptual conservatism*. This commitment is problematic for the reason that perceptual conservatism faces a range of problems, some of which we touched upon in earlier chapters.[105] As we will see, testimonial conservatives are not unaware of the potential issues here, recognising perceptual conservatism to be a problematic position from which they seek to distance their own conservative views by arguing that no such commitment obtains. I aim to demonstrate that it does obtain, presenting testimonial conservatives with a choice between embracing perceptual conservatism or abandoning testimonial conservatism altogether.

[104] Most of the material from this chapter are taken from 'The Symmetry Problem for Testimonial Conservatism' (Jope, forthcoming[c]).
[105] Insofar as perceptual conservatism is committed to entitlement theory—and unless we are to embrace scepticism it is hard to see how it could not be—chapter 2 can be seen as an argument against perceptual conservatism.

5.1. Perceptual Conservatism

One authenticity condition for forming beliefs based on perception is *that one's perceptual faculties are reliable*. If, for example, I had reasons or evidence for doubting that my perceptual faculties are a reliable guide to the truth, this would serve to defeat any justification I might otherwise have for forming perceptually based beliefs. To echo some of the discussion from the end of chapter III., merely to doubt the reliability of perception—independently of whether I have good reasons for doing so—would seem to rationally commit me to doubting that, say, what I see before me now is a laptop. I will henceforth use the label *perceptual conservatism* for the view that the acquisition of perceptual justification—that is, justification to form beliefs on the basis of perceptual experience—requires prior justification to take perception to be reliable. Likewise, *perceptual liberalism* denies this, requiring the mere absence of a defeater for perceptual reliability.

Stepping away from perception for a moment, a more general question we can toss into the liberal-conservative debating chamber is whether, given any method of belief formation M, prior reasons to take M to be reliable are necessary in order to justifiably form beliefs via M.[106] Given the array of different belief forming methods, some more general, others more specific, we can easily talk about conservative vs. liberal views regarding memory, introspection, *a priori* reasoning, universal generalisation, and testimony, to name a few.

It is clear that Wright intends his conservative framework to apply to a whole range of domains of inquiry. Indeed, each of the examples of different domains of inquiry from the previous paragraph are discussed by Wright in at least one of his various articles on this topic. However, he also notes that most of the debate thus far "has been focused on the nature of the warrant provided by a perceptual experience for beliefs about the local environment" (2012, 473). In one sense this is correct in that, as mentioned, most philosophers explicitly endorsing conservative or liberal views have tended to focus on perception—cases such as *Moore* and *Zebra* take up the lion's share of the discussion. And yet, in another sense this obscures an important fact, namely that, independently of the liberalism-conservatism debate, there is a rich body of literature elsewhere that asks more or less same questions but restricted to the domain of *testimonial* knowledge and justification. It is a curious fact that these debates have taken place independently of each other, given some clear theoretical overlap.

[106] One influential work that discusses precisely this question is Stewart Cohen's article 'Basic knowledge and the Problem of Easy Knowledge' (2002).

5.2. Testimonial Conservatism

In the literature on the epistemology of testimony one key question is whether hearers in testimonial exchanges can acquire warrant for testimonial beliefs via testimony independently of having some prior justification for trusting testifiers to be reliable. Discussion of this point has tended to be subsumed under a related discussion over whether testimony is a *fundamental* epistemic source or whether its epistemic status can be reduced to other more fundamental sources such as perception, memory, and induction. If the epistemic status of testimony can, as *reductionists* hold, be reduced to these other sources, then a form of testimonial conservatism is true: testimonial justification is acquired via a kind of inference from the speaker's utterance that P, together with reasons to trust that testimony is reliable, or that the particular testifier in question is reliable, to the conclusion that P.

This conservative, inferentialist account of the structure of testimonial justification faces similar problems to perceptual conservatism. Requiring that one first establish that the practice of relying on the word of others is a reliable way to form beliefs seems to make the acquisition of testimonial justification difficult if not impossible to achieve, requiring that subjects carry out some immense empirical investigation into the reliability of testimony as a general practice or of the general trustworthiness of an individual testifier one is confronted with, somewhat negating the practical value of relying on the word of others. This would also seem to place a heavy burden on the shoulders of the cognitively less sophisticated—for example, children—who rely on testimony to such a high degree that it seems they do not have the kind of independence of mind needed to establish testimonial reliability, nor in certain cases the necessary conceptual apparatus. And yet, there is perhaps no group that relies so heavily on testimony as a route to knowledge of the world than young children. Moreover, even setting aside these practical questions, perhaps the biggest hurdle to this kind of view is that critics point out that it is hard to see how one could establish the reliability of testimony without in some way relying on testimony, thereby inviting circularity-based sceptical worries.[107]

Those who reject the reduction of testimony to more basic epistemic sources (anti-reductionists) seem committed to a kind of *testimonial liberalism*, according to which no independent reasons to take a speaker to be a reliable, trustworthy source of testimony are strictly necessary. Liberalism about testimony takes there to be a *prima facie* kind of justification for testimonially based beliefs that subjects can enjoy merely by virtue of being a

[107] See, for example, Coady's discussion of Hume's conservative, inferentialist account of testimony and his charge that the requirement is "question-begging" (1992, 80).

hearer in a testimonial exchange. This justification can be easily defeated if, for example, there are reasons to take it that a testifier is being insincere.

While in previous chapters we have been following Wright in referring to these background supporting propositions as authenticity conditions, some in the testimonial literature such as Elizabeth Fricker use the label "validity conditions" (Fricker, 1995, 398). These include that speakers are both competent believers and sincere testifiers. Testimonial unreliability can result if either of these two conditions are not met. Testimonial conservatives such as Fricker are thus committed to the claim that prior justification for both the sincerity and competence of testifiers is necessary. As we will see below, Fricker takes this multi-dimensionality of testimonial reliability/unreliability to have important theoretical consequences.

5.3. Theoretical Symmetry

With respect to this issue of reliability of various methods, the kind of conservatism that Wright favours is an across-the-board conservatism that applies to perception, testimony, and other belief-forming methods besides. The various examples of non-transmissive inferences which he takes to have the familiar conservative structure include both perceptual cases and testimonial cases.[108] For Wright it looks like the underlying question here is whether, given any belief-forming method M, prior reasons to take M to be reliable are strictly necessary—to which the wholesale conservative answer is a resounding *yes*.[109] A possible liberal alternative is to agree about the underlying question but disagree about its answer. A liberal of this kind argues that no prior reasons to take perception to be reliable are strictly speaking necessary, and likewise no such reasons are needed for testimony. Views like these hold that our answers to these questions ought to match across different domains. I will use the term *theoretical symmetry* to talk about the kind of view that says our answers to these questions about perception and testimony ought to match—i.e. conservatism or liberalism *across the board*. Two epistemic domains A and B are symmetrical if defending conservatism (or liberalism) in A commits us to conservatism (liberalism) in B, and vice versa. Likewise, I will use the term *asymmetry* to talk about the kind of view which says that we ought to be conservative about one domain but liberal about another.

[108] Wright discusses a testimonial version of a non-transmissive inference which has the familiar conservative structure in 'Replies Part IV: Warrant Transmission and Entitlement" (2012, 461).
[109] Of course, Wright's conservatism is not restricted to these issues about reliability of method. Rather, that a particular method is reliable is merely one type of authenticity condition.

Several authors seem to go along with something like symmetry. As mentioned, Wright (2012, 2014), but also Annalisa Coliva (2015), Cecil A. J. Coady (1973, 1992), Peter Graham (2006, 2011), and Tyler Burge (1995, 1996). Some notable detractors are Elizabeth Fricker (1987, 1995) and Jennifer Lackey (2006). The following sections of this chapter will deal with a variety of arguments from Fricker and Lackey in defence of *asymmetry* (though I am introducing these terms, it is clear from how I define them that each of these three views is a version of asymmetry). Each of these three authors defends some version of conservatism about testimony, while attempting to show that this in no way commits them to a corresponding conservatism about perception.

Fricker defends a conservative account of testimony according to which knowledge can be gained through testimony only inferentially via knowledge of speakers' reliability (1995, 399). Fricker notes that corresponding conservative views in other domains are fraught with problems: "It is familiar that it is hopeless to treat knowledge through either memory or perception as inferential rather than direct" (1995, 400). What Fricker needs then is to reject symmetry—she needs argument for why we ought to treat testimony differently to memory and perception, requiring stricter epistemological standards than do these other sources. Echoing this concern, Jennifer Lackey nicely summarizes the situation:

> "If [symmetry is true], there seems to be a problem of overgeneralization here. For now it looks as though positive reasons are needed to justifiably hold, not just testimonial beliefs, but any beliefs. And this, in turn, leads us into all of the problems facing traditional internalist theories of epistemic justification, such as infinite regresses, circularity, foundations, and so on." – (2006, 176)

Both Lackey and Fricker defend versions of views which posit a default requirement on hearers to have positive reasons or evidence to trust that speakers are reliable testifiers. The corresponding view about perception is practically a non-starter, which is what these quotes are gesturing at. If we first need to establish that our perceptual faculties are reliable prior to using them to form beliefs, we cannot rely on them in doing this. But if we cannot rely on perception to establish its reliability, then it is very hard to see what other options are available—after all, it is a contingent, empirical matter whether our perceptual faculties are reliable, so *a priori* reasoning can be of no help here. Therefore, if evidence for the reliability of perception is needed in order to licence perceptual belief formation, then full-blown perceptual scepticism seems to ensue. This is why it is crucial for Lackey's and Fricker's

projects, and indeed all reductionist, inferentialist accounts of testimony, that they can establish that there are reasons to want to treat perception and testimony differently—i.e. to establish what I am calling asymmetry.

Recall that the kind of conservatism that Wright defends is not so strong as the above. Indeed, one of its key features is this thought that there is a lacuna between the recognition that one has no reasons or evidence for a proposition P and the conclusion that one thereby has no warrant to accept P. This disjunctive account of warrant thereby allows Wright to make sense of how there can be warrant to accept authenticity conditions in the absence of reasons or evidence. This suggests that a possible way for testimonial conservatives to respond is to take their cue from Wright by diluting the conservative requirements on testimonial 'validity conditions' such that it is not evidence or positive reasons that one needs in favour of the reliability of testimony, but rather a mere entitlement to trust. One could then stick to one's conservative instincts regarding testimony, embrace theoretical symmetry for both domains, while avoiding the kind of sceptical consequences of the stricter version of conservatism mentioned in the previous paragraph. Such a view would, however, face all of the familiar problems for entitlement theory raised in chapter II. I will not say anything in this chapter against the viability of this weakened form of testimonial conservatism beyond what was said against entitlement theory earlier on. The aim of this chapter is merely to show that attempts to drive a wedge between perception and testimony in order to defend conservatism about testimony without being committed to a corresponding conservatism about perception all fail. In each case, either the arguments offered in defence of asymmetry do not establish a difference between perception and testimony, or the differences they do establish are not theoretically significant. The upshot is that if one is a conservative about testimony, then one ought to be prepared to double down and commit to conservatism about perception too. To the extent that one finds perceptual conservatism—of either the stronger evidential variety or the weaker entitlement variety—problematic, this ought to cast doubt over conservatism about testimony.

5.4. Neutrality and Non-Neutrality

There are clear differences between knowledge through perception and knowledge through testimony. This much is obvious. What is less obvious is whether these differences give rise to different epistemological standards. How might we begin to think about this? I think that a reasonable starting point should be a presumption that the standards are the same unless there are good reasons to think otherwise. After all, one can ask the very general question of whether,

given any epistemic method or source, prior reasons to take the method to be reliable are necessary. The simplest kind of answers to this question will either be that they are necessary or that they are not necessary. If the answer is that in fact things are more complicated and it all depends on which source we are talking about, then we need to be persuaded of that. This places something of a burden on the asymmetrist to show why we ought to think there are theoretically significant differences between perception and testimony.

Fricker's case for applying stronger standards to testimony begins with the observation that it is part of the nature of a perceptual experience "which is a state in which the state of affairs experienced as instantiated is represented to the subject as obtaining" (1987, 74). Or perhaps more cautiously, where an experience is an experience as of a certain state of affairs, the state of affairs is represented as obtaining.[110] So a perceptual experience as of, say, a cup on the table in front of one is an experience in which there is a representation as of a cup on the table and, crucially, this representation is represented *as true*. This direct, immediate link between experience and perceptual representation means that such experiences cause a direct impulse to believe: "the very nature of an experience means that it is a state with an internal, unmediated impulse to believe in what is experienced" (*ibid*). And thus beliefs based on perceptual experiences need not be based on any kind of inference. However, Fricker thinks this is not the case with testimony. "There is nothing", she argues, "in a hearer's perception of an utterance as, say, an assertion that P which is, from either a phenomenological or an epistemological point of view, his *experiencing P itself as instantiated*" (*ibid*). This is the crucial difference we were looking for. Fricker locates the asymmetry between perception and testimony in the thought that perceptual experiences are representations of states of affairs *as true*, while assertions are not. And thus, hearers need to perform a kind of inductive or abductive inference, from the speaker's utterance that P, to the conclusion that P, via some additional premise to the effect that the speaker's testimony can be relied upon.

Granting Fricker's point that there is something in the nature of a perceptual experience that the state of affairs represented in the experience is represented as true, why should we think this is not also the case with testimony? To help get some clarity on the issue, let's first consider a kind of representation of a state of affairs in which it is trivially not the case that the state of affairs is represented as true. Consider a painting of a particular scene in which some event is taking place. The event is represented in the painting by the artist, but is it represented *as having happened*? We surely can imagine a kind of case in which it is—perhaps an artist

[110] Leaving room for the possibility of certain non-representational experiences.

working for the police has painted a representation of an event based on witness testimony. But this is beside the point. The question is whether it is part of the very nature of representational painting in general that this the case. Clearly the answer here seems to be no. Knowing nothing else about a picture, about why it was painted and by whom, we would feel no compulsion to believe that the events depicted had actually taken place, nor would there be any expectation of believing. All else equal, a painting of an event taking place is neutral with respect to the issue of whether the event did in fact take place. With this contrast case in mind, we are better positioned to ask the question of whether testimony is more similar to perception, in which the represented state of affairs is represented as true, or more like a painting, in which the state of affairs is not represented as true. I think it should be relatively clear that an assertion is more similar to a perception than a painting in this regard. Consider the fact that seeing or hearing about something surprising might cause us to update our picture of the world or else respond with incredulity. If someone tells me that the president of France has been awarded the Olympic gold medal in the one hundred metre sprint, I will probably respond with surprise and incredulity. If someone shows me a painting of the same, my response will be more likely one of mild amusement at the artwork. It would not normally be proper of me to respond to the painting as if my friend had just made the corresponding assertion about the president by asking for evidence to back it up. An assertion that P is not neutral with respect to the truth of P in the same way that something like a painting is.

Reflection on the nature of the assertoric speech act reveals some interesting characteristics that go some way to explain why we are inclined to take assertions to be representations of states of affairs *as true*. Analyses of assertion usually point to two essential characteristics: content and force. Assertions have content in the sense that they represent propositions. But this is true of other types of speech acts as well: questions, suggestions, guesses. What distinguishes assertions from these other speech acts is *assertoric force*. Assertoric force is more than simply representing a proposition as true via a particular speech act. It is also to represent oneself as having the authority to make an assertion, and thus as having met the requisite norms of assertion, with the intention that one's audience themselves form the corresponding belief. Far from being neutral with respect to the truth of the represented content, an assertion is a speech act that is defined by its non-neutrality.

One further point against the thesis that assertions are truth-neutral is that comprehension of assertoric content cannot occur in a truth-neutral manner. That is, the inferentialist picture according to which we first comprehend the content of an assertion in a neutral manner before then choosing whether to infer its truth is neither theoretically nor

empirically well founded. On the theoretical side, Patrick Rysiew argues that hearers cannot comprehend the content of assertions unless they presume the speaker is being truthful (2012). This entails that arriving at comprehension of assertion in a truth-neutral manner in the way proposed by the inferentialist is impossible. I cannot interpret your assertion *as an assertion* without a presumption that you are attempting to be truthful, for otherwise the verbal act underdetermines whether you are asserting as opposed to guessing, joking, acting, and so on. Furthermore, even within the scope of assertoric speech, given loose ways of talking, the verbal act underdetermines which of a number of semantic interpretations is the correct one. The presumption of truthfulness is necessary in order for me to understand that when you say *everyone is coming to dinner tonight*, you do not mean this to be taken literally. In order to get to the correct interpretation of your assertion, I must presume that you are adhering to the proper rules and conventions of linguistic communication such as the Gricean *Maxim of Quality*—'try to make your contribution one that is true' (1989, 27). On the basis of considerations such as these, Rysiew concludes that "if the presumption of truthfulness is required just to arrive at what a speaker is telling, questions about whether/why one should believe what one's told can only arise against the background of the presumed reliability of testimony" (2012, 293). A presumption of truthfulness is thus the default, which means that neutral comprehension of testimony that is prior to truth assessment is an inferentialist illusion. This presumption should not be conflated with the conservative claim that we need prior justification to presuppose that speakers are reliable. The former is a claim about the psychology of understanding, the latter is a claim about the structure of justification. The takeaway here is simply that assertions are not truth neutral because understanding them requires the presumption of truth. That is compatible with the liberal claim that justification to believe them requires no prior justification to take speakers to be reliable. Taking speakers at their word may be a matter of basic entitlement.[111]

On the empirical side of things, the psychological literature on *truth bias*, sometimes also called *truth default theory*, corroborates the foregoing theoretical insights. For example, Daniel T. Gilbert and his team provide strong evidence for thinking that acceptance of communication coincides with comprehension (1990, 1991, 1993). That is, subjects do not first comprehend an idea that is presented in communication before then deciding whether to accept or reject it. Rather, acceptance is the default setting and only after having accepted a communication as true does the possibility of rejection as false arise. Gilbert labels this the

[111] See Burge (1995) and also Graham (2018[b]).

Spinoza hypothesis, on the grounds that Spinoza held the view that "all ideas are accepted (i.e., represented in the mind as true) prior to a rational analysis of their veracity, and that some ideas are subsequently unaccepted (i.e., represented as false)" (Gilbert et al, 1990, 601). The evidence for this Spinozan view is that processing a proposition as false requires more time and cognitive capacity than processing as true. Moreover, they find that this is no mere social convention but is part of the very nature of the cognitive mechanisms in play. Incredibly, even if subjects are primed to *expect* false information, still it looks as though there is an initial representation-as-true which precedes a subsequent judgement of falsity:

> "In short, knowing ahead of time that information would be false apparently did not enable subjects to adopt a skeptic's set and represent the information initially as false. This suggests that the initial coding of ideas as true may (as Spinoza suggested) be an operation that is not readily amenable to voluntary control." (Gilbert et al, 1990, 607).

What the Spinozan hypothesis means for the current discussion is that assertions, by default, are experienced as true by hearers, according to the literature from psychology on truth bias. In conjunction with the claim above that force as well as content is a distinguishing feature of assertions, this provides a very strong case for rejecting the claim that assertions are not representations of states of affairs *as true*. The truth-neutrality account of assertions is neither a plausible account of the nature of the content of assertions, nor a plausible account of the actual cognitive mechanisms underpinning our receipt of testimony. We will need to look elsewhere for a theoretically important difference between perception and testimony.

5.5. Degrees of Reliability

A thought that both Lackey and Fricker pick up on is the idea that testimony just is *de facto* more unreliable than perception and thus more scrutiny is required of us when forming testimonial beliefs as compared to perceptual. In considering the liberal policy of believing in accordance with speaker's utterances, absent positive reasons to take the speaker to be unreliable, Fricker writes that "the proportion of utterances which are made by speakers who are either insincere or incompetent is far too high for this to be an attractive policy" (1987, 76). Let us grant the point that testimonial unreliability is significantly more unreliable than perception, is this difference epistemologically significant? Suppose that it is. Peter Graham points out that this kind of argument presupposes a problematic principle according to which

differences in degree (of reliability) determine differences in epistemic kind (Graham, 2006, 101). Arguably, perception is less reliable than introspection, memory less reliable than perception and testimony less reliable than memory. Degrees of reliability exist between each of these epistemic sources. Thus merely pointing out a difference in degree of reliability is not enough to establish asymmetry because there is a whole spectrum of degrees of reliability along which to place a variety of epistemic sources but only a single line to be drawn between conservatism and liberalism. What testimonial conservatives need to argue for in order to establish the asymmetry claim is that the difference in reliability between perception and testimony is theoretically significant. Simply pointing out that testimony is generally less reliable than perception as a way to establish asymmetry begs the question since it remains to be established that the theoretically significant threshold the asymmetrist needs is to be located somewhere along the spectrum between the precise loci of perception and testimony. Suppose we grant the thought that there is some threshold along a spectrum of degrees of reliability of various epistemological sources and methods. This is enough to establish that there is asymmetry between *some* methods. It nonetheless remains to be established just where exactly the threshold lies and between which methods. For consider the possibility that perception falls below the threshold. Perception is somewhat unreliable, testimony more so, but both fall below the threshold and thus symmetry is true of these two methods. What possible argument could the asymmetrist who wants to locate the threshold between testimony and perception give for excluding this option? A third possibility yet is that the threshold is in fact much lower, to the extent that both perception and testimony pass the bar needed to be treated liberally. Once again, the asymmetrist who wishes to ground their theory in differences in degrees of reliability needs some independently plausible argument for excluding this version of the threshold view, and yet it is very difficult to see on what grounds one could attempt such an argument.

Lackey offers a related but nonetheless distinct argument for symmetry on the grounds of a difference in reliability. Her argument is a modal one, appealing to the familiar thought that the possible worlds in which most of our perceptual beliefs are false are quite distant from the actual world—the familiar brain-in-a-vat scenario, for example. This means that perceptual reliability, though not perfect, is relatively modally robust. Compare this with testimonial unreliability. Lackey thinks that worlds in which most of our testimonial beliefs are false are much closer: “indeed, for many people this is true in the *actual* world” (2006, 177). The kinds of examples she has in mind here are scenarios in which one is raised by parents who are in a cult, or in societies that are highly superstitious or under governments that are highly corrupt. What she takes this to show is that the chance of testimonial error is much higher than

perceptual error and for this reason "the rational acceptance of the reports of others requires positive reasons in a way that is not paralleled with other cognitive faculties" (*ibid*.). Clearly this argument again appeals to the differences-of-degree-equal-differences-in-kind principle that we have rejected. However, there are some further worries for Lackey's argument that, upon reflection, seem to point us away from the kind of asymmetry she is trying to motivate.

Radical sceptical scenarios are easy to posit. The most familiar ones involve radical perceptual deception such as the brain-in-a-vat, evil demon or Matrix scenarios. Though we ordinarily take such scenarios to be modally very distant from the actual world, the sceptical bite comes from the acknowledgement that they are perfectly logically consistent with our subjective experiences. Furthermore, we might easily imagine equivalent sceptical scenarios for other sources of knowledge besides perception. Consider Bertrand Russell's argument that it is logically compatible with all of our apparent memories of the past that, in fact, the world simply sprang into being mere moments ago (2005, 132). Just like the perceptual-deception scenarios, Russell's memorial sceptical scenario would seem to be perfectly internally consistent and consistent with our current subjective experiences. It is important to bear in mind that in these radical sceptical scenarios, virtually all of our perceptual (or memorial) beliefs would be false. With this in mind, it is far less clear that there are real-world scenarios of radical *testimonial* unreliability. Consider the examples Lackey appeals to, cults and oppressive regimes. Do these examples merit the label of *radical* deception scenarios, on a par with the brain-in-a-vat and Russellian scenarios? It would seem that they do not. After all, even living in a cult, while one could easily have many false testimonially-based beliefs (beliefs about the world outside the cult, beliefs about the origin of humanity, moral beliefs) one presumably has lots of mundane, ordinary, true testimonially-based beliefs also (beliefs about what time it is, about what is being cooked for dinner, about the location of objects in one's environment). While it is perhaps part of the very essence of cults that they entail some kind of testimonial unreliability, the beliefs in question are typically restricted to a certain domain (religious, political, moral). It is therefore not clear that these scenarios are at all on a par with the familiar radical (perceptual, memorial) sceptical scenarios. What Lackey needs to substantiate her claim are cases in which practically *all* of one's testimonially-based beliefs are false and then some reasons for thinking such scenarios are modally closer to the actual world than the perceptual cases. Far from this being an achievable goal I think there are reasons for thinking that things are in fact the other way round: the perceptual sceptical cases are modally closer than the radical testimonial sceptical scenarios.

What would it mean for all of one's testimonial beliefs to be false? Conceiving of such a scenario is not as simple as in the perceptual case. Firstly, in the perceptual scenario there is no need to think that radical perceptual unreliability need prevent one's perceptual beliefs being generally coherent. However, in the testimonial case things are not so clear cut. Consider a subject—let's call him Truman—who is in the unfortunate scenario of being lied to by those around him. It is perfectly conceivable that those in Truman's life seek to deceive him through false testimony. Truman is told all sorts of lies about the name of the small town he lives in, the world beyond the town, and the lives of those he lives with. It's easy to conceive of a scenario of massive orchestrated testimonial deception. But this doesn't put testimony on a par with perception. For that to work, we need to conceive of a scenario of not just massive deception but total, radical deception. And the obstacle we encounter in attempting to do so is that, while in the radical perceptual case, perceptual unreliability need not affect reliability in other domains—the envatted subject may well have perfectly excellent powers of reasoning and she may also have impeccable memory. Things are different for testimony because, unlike for perception and memory, it is very easy to compare testimonial evidence with other kinds of evidence. Consider what will happen when Truman ask simple questions about objects in the immediate environment that can easily be checked by other means such as 'What's that you've got in your hand?' or 'Does the cheeseburger come with fries?' Or consider what will happen when Truman asks questions about the recent past or near future the answers to which will soon become apparent to him anyway such as 'What are we having for dinner?' or 'Did it rain last night?' The point is that in order for the massive, orchestrated deception to function, a certain amount of the testimony that Truman receives needs to be reliable. If it isn't, it will very quickly become obvious to him that this is the case and if he is a normal functioning person, he will surely stop believing anything he is told.

A case in which testimony is *globally* unreliable is much harder to conceive of than our Truman case. In fact, there are reasons for thinking that such cases could not possibly exist. The reason for this is that the practice of testimony itself will not survive wholesale unreliability. Suppose that we found ourselves in a world in which the propositions that others express when they seem to be making assertions bore no relation to the truth. The unreliability of these utterances will quickly become obvious: when it is raining, people will declare that it is not; when it is night, people will declare that it is day. One consequence of this is that no one would believe such barefaced lies. Moreover, as C. A. J. Coady argues, in a world in which there were no correlation between apparent reports and the facts, it is hard to see why we should think there are reports at all (1992, 84). Coady imagines encountering a Martian community

whose language we could translate and who seem to systematically say things that we and they can observe to be false. In such a situation, he wonders, "what reason would there be for believing that they even had the practice of reporting?" (1992, 84). Peter Graham concurs with Coady that if all assertions are falsehoods, then the practice of asserting will come to an end because hearers will have no reason to accept assertions, giving speakers no reason to make them (2000[a], 698).[112] This shows why the radical sceptical testimonial cases, far from being modally closer to the actual world as Lackey claims, are arguably not even coherent scenarios. One might wonder whether we can nonetheless conceive of sceptical testimonial scenarios in which, if not all, then *most* of the assertions made are false. But Graham goes on to deny even the coherence of this less radically unreliable scenario. Even if, say, the imagined Martian community make equal amounts of true and false reports, still the practice of assertion would not survive because relying on the word of others would be no better than flipping a coin (2000[a], 699). Unless speakers mostly utter truths, the economy of testimony will thus break down.

The sustained practice of testimony conceptually implies a certain degree of testimonial reliability without which members of a community would have no reason for engaging in the practice. This is not yet to say that radical sceptical testimonial scenarios are impossible. Rather, it is an empirical claim about the institution of testimony, that wholesale unreliability would mean hearers have no reason to believe what speakers say, which in turn would mean speakers have no reason to give testimony. This argument is closely connected to Rysiew's argument, discussed in the previous section, that the practice of testimony requires a presumption of truthfulness in order to function (2012). And since wholesale unreliability would lead hearers to forgo the presumption of truthfulness, future attempts to assert would fall flat, no longer interpreted as truth-apt declarative utterances. The consequences of this are far more radical than that the economy of testimony would likely break down. The point here is that the utterances such speakers make would no longer count as assertions because they would carry no assertoric force. If we accept the plausible claim that assertoric force is a constitutive component of assertions, then any speech act that ceases completely to bear any degree of force will thereby cease to even count as an assertion. That is why radical sceptical testimonial scenarios, far from being modally close, are not even coherent possibilities.

Reflection on differences in degrees of reliability does not therefore seem to offer a way to establish the kind of asymmetry claim that the testimonial conservative seeks. We must

[112] Graham draws on this point again in later work (2010, 2019).

again look elsewhere for our theoretically significant differences between perception and testimony.

5.6. Unpredictability of Testimony

There are two dimensions to testimonial unreliability: incompetence and insincerity. A testifier can provide unreliable testimony if they are themselves incompetent believers. But a testifier who is not an incompetent believer can still provide unreliable testimony by being insincere. These two dimensions to testimonial unreliability make testimony more unpredictable than perception, according to Lackey, because perceptual unreliability, by comparison, is one-dimensional (2006, 176).

This argument is somewhat similar to the differences-in-degree-equal-differences-in-kind argument from above—both arguments derive higher epistemic standards for testimony on the grounds of greater testimonial unreliability. But whereas above we weren't able to find principled grounds to set an unreliability threshold—beyond which the higher epistemic standards kick in—this argument promises to provide such grounds. Contrary to the earlier argument, it isn't the *degree* of unreliability that matters, but the complexity, or number of *dimensions* that are important. Where there is more than one dimension to the possibility of unreliable bases for belief, subjects need to be on guard against such possibilities.

Once again, this argument fails to establish the asymmetry. To see why, assume that the principle it relies on is correct, namely the following. Where there is just one dimension to cognitive unreliability liberalism is true, but where there is more than one dimension to cognitive unreliability conservatism is true. It is not clear what the argument for this principle is meant to be, but it at least feels relatively less arbitrary than identifying some particular degree of reliability as the threshold beyond which conservatism kicks in. However, even with this principle in play, it still does not force the wedge between testimony and perception that the asymmetrist needs.

The first thing to note is that to suggest that perception is one-dimensional in this way is an oversimplification. Perception is multi-modal in that it may represent the size of an object, its shape, distance, motion, relation to oneself, and so on. These multi-modal aspects of perceptual representation may come together to give a subject a cohesive picture, but they rely on a variety of *perceptual modalities*, or *sensory streams*.[113] A reliable perceptual system depends on each of these many modalities functioning reliably and on their coming together in

[113] Thanks to an anonymous reviewer for suggesting this point.

a reliable way. Visual agnosia occurs when one of these sensory streams breaks down, giving rise to an inability to process sensory information such as distinguishing objects from their backgrounds or motion detection. Perceptual representation is thus multimodal and the many dimensions to perceptual unreliability correspond to the many forms of visual agnosia.

Despite this recognition of perceptual multimodality, we might nonetheless wonder whether there is something unique about the *agential* nature of testimony that would support asymmetry. The thought would be that since speakers are free to choose to deceive us in a way that our perceptual systems are not, we ought to be on our guard against this possibility and that is why conservatism is true of testimony but liberalism is true of non-agential sources like perception. There are two assumptions here that we might reject. Firstly, Graham argues that the agential nature of testimony does not give rise to the kinds of strict epistemic standards that conservatives say it does (2000[b], 2004). If Graham is right about this, then differences in whether sources involve the possibility of deception or not will have no theoretical significance for the conservatism-liberalism discussion. I think Graham is indeed right about this. However, I want to push back against the second assumption. This assumption is that deception is unique to testimony. Assuming for the sake of argument that the possibility of deception does give rise to conservative justificatory standards for testimony, why should we assume does this not translate into a corresponding argument for perception? The obvious answer is that speakers have agency and this enables them to deceive us while our perceptual faculties do not have agency and so are unable to deceive us. Perhaps the multimodality of perception is beside the point if the dimension of *insincerity* in testimony has no perceptual counterpart for perception. This does not, however, seem to be the case. Consider the fact that con-artists, pranksters, illusionists and other practitioners of the deceptive arts may seek to cause us to form false beliefs based on perception in all sorts of ways. I visit a Broadway magic show and watch the magician enter a cupboard door only to have vanished into thin air a moment later. It is all an illusion of course, and the magician is preying on my credulity, but it is nonetheless an act of deception on a par with insincere testimony. In cases such as this, even though my perceptual faculties themselves are not malfunctioning, illusionists and other tricksters can lead me to form false perceptual beliefs through deceptive means. Deception is not unique to testimony.

Just like perception and testimony, we can also think about memorial unreliability along multiple dimensions. In particular, we can perfectly well think about different dimensions of memorial unreliability that mirror those discussed by Lackey. The standard way for memory to mislead is a simple act of misremembering. I seem to remember taking my keys with me when I left the house, but it turns out I left them at home and have misremembered. In such a

case, my memory is simply letting me down, but it is not being 'deceptive' in the agential sense of intending to lead me astray. In this sense, it is on a par with the dimension of 'incompetence' that Lackey discusses in relation to testimony. Contrast this with a case involving the intriguing phenomenon of memory implantation.[114] In memory implantation cases, subjects in controlled experiments have been led to form false memorial beliefs by experimenters manipulating them using a variety of different means. In one experiment, subjects are shown photographs of their childhood that have been manipulated in such a way to depict an event that did not happen as having happened. In another experiment, subjects are given a narrative of events that happened to them as a child with one false event mixed up among several other true events. Unaware that a false event has been slipped into the narrative, a significant number of subjects adopt the false event as part of their memory of the series of events. What cases like these show is that it is in fact relatively easy for agents to manipulate us into seeming to remember things that did not occur. There is thus no reason to think that testimony is special in the sense that it is only through testimony that we can be deceived by agents who seek to deceive us. Deception can take many forms. If it is the multidimensionality of testimonial unreliability that gives rise to testimonial conservatism, then the fact that there are multiple analogous dimensions to perceptual and memorial unreliability also seems to commit conservatives to perceptual and memorial conservatism.

5.7. Perceptual Homogeneity, Testimonial Heterogeneity

The next argument for asymmetry that we will consider comes again from Lackey. Her thought is that an epistemologically important difference between perception and testimony is that perceptual sources are homogenous while testimonial sources are heterogenous (2006, 177). To see what she means by this, it is best to run through the two contrast cases she presents.

> *Perceptual Amnesia*
>
> Suppose that after her involvement in a car accident, Olivia has complete amnesia with respect to her perceptual faculties, that is, she remembers nothing about either the workings or the deliverances of such faculties. After leaving the hospital, she stops at the store to buy some groceries, bumps into some acquaintances on her way home,

[114] For discussions of this phenomenon see Loftus et al. (1995), Wade et al. (2001) and Lindsay et al. (2004).

watches an episode of *Seinfeld* on TV while eating dinner, and spends some time on the internet before going to bed. Along the way, Olivia forms perceptual beliefs about all sorts of things, including beliefs about the vegetarian items that Trader Joe's carries, the kinds of trees losing their leaves, the number of children her acquaintance now has, which *Seinfeld* episode is on, and the colour of the background of the MSN website. Now, because of her perceptual amnesia, Olivia's acquisition of these perceptual beliefs is not governed by any acquired principles of perceptual belief formation. But even in the absence of such principles, it seems reasonable to conclude that the overall status of Olivia's daily perceptual beliefs would be very high epistemically.

Testimonial Amnesia

Edna, Olivia's best friend, was in the same car accident that caused Olivia's perceptual amnesia. In Edna, however, the accident caused *testimonial* amnesia: she remembers nothing about either the workings or the deliverances of testimony. After leaving the hospital, Edna's day was nearly identical to Olivia's. For instance, she stopped at the same grocery store, bumped into the same acquaintances on the way home, watched the same episode of Seinfeld, and visited the same internet sites before going to bed. Now, because of her testimonial amnesia, Edna's acquisition of testimonial beliefs along the way was not governed by any principles of testimonial belief formation. As a result, Edna trusted to the same extent all of the testimonial sources she encountered throughout the day—which included a copy of the National Enquirer that she read at the grocery store, her acquaintance's 3-year-old daughter, the characters of Jerry and George on Seinfeld, and an extremist, evangelical Christian internet site she stumbled upon while surfing the web—and she believed everything that she was either told or read along the way—which included testimony that a woman from Georgia was abducted by aliens, that there are real princes and princesses at Disneyland, that licking the envelopes of cheap wedding invitations can lead to one's death, and that those who are gay will be sent to eternal damnation.

The intuition we are supposed to have towards these cases is that Oliva's perceptually based beliefs are justified while Edna's testimonially based beliefs are not justified. The explanation that Lackey offers for this difference is that perception is homogenous—there is not much difference between Olivia's various perceptual beliefs and the ways they were formed. Her

belief that there are leaves on the tree was formed by looking at the leaves on the tree. Her belief about the kinds of vegetarian items on offer in the supermarket was formed by looking, and so on. Because of this homogeneity, Lackey argues that "subjects do not need to be very discriminating in order to be reliably in touch with the truth" (2006, 177). On the other hand, the sources of testimony are a heterogenous group, with some sources being less reliable than others. According to Lackey it is the fact that Edna's amnesia renders her insensitive to the differences between testimony's various sources that lead her astray. She trusted sources that she wasn't aware she shouldn't have trusted, such as the New York Post and the 3-year old's depiction of Disneyland. Importantly, in both cases, were we to stipulate that our subjects did *not* suffer from amnesia, Edna's beliefs would have been significantly different while Olivia's would not. According to Lackey, what this demonstrates is that possessing acquired principles governing the acceptance of testimony is necessary for testimonial justification, while the same is not true for perception.

Has Lackey identified an epistemically important difference between perception and testimony here? There are good reasons to think not. Consider what testimonial conservatives want to say about young children who are incapable of meeting the conservative requirement on testimonial justification. It is plausible that even children who lack concepts such as sincerity, lying, reliability and so on can know things via testimony. Indeed, young children would seem to rely on testimony as a source of knowledge about the world to a greater degree than do adults. And yet, conservatism threatens to make it very hard for such children to properly acquire knowledge and justification via testimony given their supposed lack of the relevant concepts and lack of sophisticated abilities to make judgements about sincerity, competency, and so on. Reductionists of course have things to say about this objection. For example, Fricker attempts to sidestep this issue by making an exception for young children, arguing that they do not need justification to trust in the reliability of testimony while they are still in the developmental stage (1995, 403). But as they grow up and they acquire discriminative capacities and knowledge of what people and institutions are like, the conservative constraints kick in and they need to engage those discriminative capacities and make use of that background knowledge. Suppose that, as is highly plausible, young children can acquire knowledge via testimony even though they lack the kind of background supporting reasons that reductionism generally requires of hearers in testimonial exchanges.[115] This raises

[115] On the possibility that children cannot, in fact, know or justifiably believe on the basis of testimony see Van Cleve (2006, 68). I concur with Graham that denying that children can know things they are told by parents and teachers is wholly "implausible" (2018, 3024).

the question of why we can't simply treat Edna like a child. Edna has lost the necessary concepts such as testimonial unreliability, lying, and so on, rendering her unable to discriminate between trustworthy and untrustworthy sources of testimony, much like a young child. Edna's testimonial beliefs are therefore unjustified only if young children have no justified testimonially based beliefs. But young children plausibly do have some justified testimonially based beliefs and therefore so does Edna.

A further issue with these cases is that some of the differences that Lackey is picking up on seem to be merely a result of differences in the two narratives she tells. It is just a feature of the stories that Olivia forms lots of true beliefs while Edna forms lot of false beliefs, but we can easily imagine different stories, equally as plausible, that give us different results. For example, suppose that, as she is leaving the hospital Olivia believes her path ahead is clear but as she steps out she walks straight into a glass door. She then wanders past a pond and notices a pole that is sticking up out of the water. She believes the pole is bent, as it appears to be bent just at the point where it hits the water, though in fact the stick is straight, and this is a common illusion caused by refraction of light. Olivia then comes across a skilfully drawn pavement sketch artist drawing a perspectival illusion, intended to appear as a hole in the ground when viewed from the correct angle, and forms the corresponding false belief that she is in danger of falling through. Olivia next sees a billboard poster which displays the Müller-Lyer illusion, forming the corresponding belief that the two lines are off different lengths. The point here is that Lackey simply builds into her narrative that Olivia forms lots of true beliefs via perception despite her amnesia and concludes that the epistemic status of those beliefs is unaffected by the amnesia. But what these examples illustrate is that Olivia will be led to form false beliefs given that she is not on guard against these ordinary, everyday visual illusions.

In addition to leading her to form false beliefs based on perception, Olivia's amnesia will also greatly hinder her ability to form any judgements at all for a wide class of perceptually-based beliefs. Having no understanding about the workings and deliverances of perception, no memory of past perceptual experiences, no awareness of how information or conceptual content can be encoded through visual representation, and no understanding of how different perceptual modalities represent information, Olivia will be incapable of taking in lots of the information that would otherwise be available to her. She may in some sense 'see' a sign telling her not to cross the road, but she cannot see *that* the sign is telling her not to cross the road. For that would require her to understand that perceptual experiences can encode instructional information, to remember which shapes correspond to which words and concepts, and so on. She may see an oak tree in the garden outside the hospital, but she does not see *that*

it is an oak tree. She may appreciate the structure of the tree but she does not recognise it *as an oak tree*. The content of perceptual experience is in part a function of what we expect to see, of our background beliefs, and of the concepts we possess (Coady, 1992, 147). Robbed of these concepts, beliefs and expectations, Olivia would not be able to judge that the items on the shelf in Trader Joe's are vegetarian or that there is a particular episode of *Seinfeld* on the television. Of course, the same point applies to Edna about whom we ought to say she would not, in fact, be able to form beliefs based on testimony since a complete and total lack of understanding of the practice of testimony and absence of memories of testimonial interactions would render her incapable of recognising an assertion *as an assertion*. Moreover, we might reasonably think that Edna's testimonial amnesia will affect not just her ability to form testimonially based beliefs but many ordinary perceptual beliefs too. The vast array of concepts that we use to make sense of experience are learned via testimony, meaning that Edna's testimonial amnesia will rob her of these concepts, rendering her unable to make sense of her experiences. Consider that in Lackey's description of the case, Edna is able to recognise and make use of a newspaper. But the concept of a newspaper is not represented by her visual system. What is represented by the visual system is the size, shape, location, colour, etc., of the object. In order for her to recognise that what she sees is a newspaper—which is necessary in order for her to form beliefs based on reading the newspaper—Edna needs to deploy the concept of a newspaper. However, since we acquire such concepts via testimony, and since Edna's amnesia means she has lost all memory of the past deliverances of testimony, she will presumably no longer have the concept of a newspaper. Likewise for the concepts needed to form beliefs about trees, chairs, tables, and so on, each of which is acquired via testimony. This suggests that Edna's unfortunate situation is indeed, as suggested above, much more like that of a child; a very young child who has yet to acquire the concepts needed to interpret her experiences.

The general lesson here is that these amnesia cases are vastly under-described. Exactly which beliefs a subject would and would not be able to form via perception or via testimony in the absence of any knowledge about the workings and deliverances of those sources or memories of past executions of them is a question that would require a much more detailed and empirically-informed discussion.

5.9. Concluding Remarks

We have seen a number of arguments that seek to establish an epistemically important difference between perception and testimony. Those arguments either fail to establish a difference or they establish some difference but not one that has the kind of epistemological significance needed to support asymmetry. A negative conclusion can be drawn in support of symmetry. None of the asymmetry arguments succeed, therefore the default position is that we ought to put perception and testimony on a par. Why should this be the default position? Because the very general question we can ask of any belief forming method M is whether positive reasons to trust in the reliability of M are necessary for acquiring knowledge and justification via M. Absent reasons to think that we ought to give different answers to this question depending on which kind of method or source we are talking about, the default position should be to give the same answer across the board. Either we take the conservative line and demand that positive reasons to trust are necessary for perception, testimony, memory and so on, or we take the liberal line and deny such reasons are necessary, or we take a mixed approach. But in the latter case we need argument for why we should treat different sources or methods differently.

6.

6. Conclusion

The main concern of this thesis has been with a family of issues grouped together under the heading *The Structure of Warrant*. The main goals of the thesis have been to identify some novel problems with the conservative, internalist, entitlement-theoretic views, and to explore the options for a liberal, externalist, neo-Moorean alternative. The resulting view is perhaps going to be not ultimately very satisfying to readers with prior reservations about externalist, or neo-Moorean, theories in general. It has not been my goal to settle these old disputes. It has, however, been my goal to demonstrate that the view can accommodate intuitions of transmission failure in at least the most clear-cut cases and is able to respond to charges that such a view would be incoherent.

In the first two chapters I presented a number of problems for the orthodox views on the market. In chapter 1. I discussed one of the supposed advantages of the transmission failure diagnosis, namely the idea that it presents a way to safeguard epistemic closure principles while accounting for some of the intuitions which led to their rejection. I argued there that in distinguishing between transmission and closure we need to be cognizant of the fact that discussions of epistemic closure in more recent years tend to formulate closure principles diachronically, along the lines of what Williamson and others call *intuitive closure*. The problem I highlighted there was that once we do this, it becomes much harder to maintain that we can safeguard closure by rejecting transmission. In order to substantiate this claim I distinguished between closure and transmission principles for both propositional justification and doxastic justification, arguing that it is difficult to see how closure and transmission failure for doxastic justification could come apart in the way needed for the proposed safeguarding of closure. Setting aside this concern, I then went on in the second chapter to explore a novel problem for the kind of view that says transmission fails in the relevant cases because the conclusions at issue in the problematic inferences enjoy a unique kind of status in the structure of justification. This kind of view holds that the propositions in question are warranted by default—that is, we are rationally entitled to accept them independently of any cognitive accomplishment. Entitlement theory faces up to a number of well-known issues, in response to which entitlement theorists have things to say. I argued that even granting entitlement theorists these responses, a residual question remains regarding the apposite degree of confidence licensed by an entitlement. I broke answers to this question down into two main categories, presenting the situation as a dilemma by arguing that neither category of answer is sustainable.

The upshot of the dilemma is that either entitlement theory is not well motivated or it leads to irrational doxastic attitudes.

In light of the dilemma for entitlement theory and the problem facing the transmission failure diagnosis for doxastic justification, I then went on in chapter 3 to suggest that it is worth exploring what other options there are for a theory of transmission failure that could capture intuitions in the relevant cases. I noted that it is often assumed that externalist theories of justification and knowledge have a particularly hard time in accounting for transmission failure but that we can resist this assumption by first clarifying which kinds of inferences are clear-cut cases of transmission failure and which are more controversial. Having made these classifications, I then went on to discuss two extant attempts to offer an externalist framework for understanding transmission failure. I argued that the resulting views face a number of issues including not being well suited to accommodate intuitions about the most clear-cut cases. I then explored ways for reliabilist and safety theory accounts of justification and knowledge to account for transmission failure. The resulting views are ones that offer a general account of transmission failure in terms of what I called positive and transparent epistemic dependence. Though these views get the correct prediction in the clear-cut cases, they do, however, predict the Moorean inference is transmissive. This raises the question of what to say about the initial suggestion that *something* about the Moorean inference seems epistemically fishy. In order to address this issue I appealed to a familiar notion that although the inference does not exhibit positive or transparent dependence, it does exhibit negative dependence, one consequence of which is that merely harbouring doubts about its conclusion will be enough to rationally commit us to doubting its premise.

The resulting view was then defended against two potential challenges in chapter 4. The first challenge consisted of the claim that externalist responses to scepticism are inherently unstable. The instability argument alleges that if we pursue the implications of externalism sufficiently far, we find ourselves forced to accept the very premise we set out to reject in responding to scepticism. I granted the devastating implications of this conclusion for externalism if the argument that led to it could be sustained, however I then showed that the externalist need not accept the conclusion on the grounds that a careful analysis of the instability argument reveals hidden internalist assumptions about evidence. I showed that once we clarify the commitments of an externalist account of evidence, we find that the alleged instability disappears. The second challenge I addressed came from a different direction, namely the argument that Moorean-style reasoning commits a kind of category mistake. The so-called non-belief version of hinge epistemology holds that having a rational basis for belief

in the conclusion of a Moorean inference is simply an incoherent ideal. This argument was then scrutinised, and it was shown that it appeals to a principle about rational basing or grounding that is in conflict with some fundamental rational commitments.

Finally, in chapter 5. of the thesis I argue that conservatism about testimony entails a commitment to conservatism about perception. Much of the discussion in the earlier chapters was dedicated to offering a critique of conservative accounts of the structure of justification. Interestingly, although these accounts ought to apply to a wide range of different epistemic domains, the literature has tended to focus on cases involving perceptually based beliefs, as is of course the case with the focus on the Moorean inference. I noted that it is an intriguing fact the debate around conservative and liberal approaches in the literature on the structure of justification has taken place independently of a different body of literature that discusses more or less the same issue but restricted to the domain of testimonially based beliefs. I characterised a certain view in the literature on testimony as testimonial conservatism and noted that defenders of this view often claim not to be committed to corresponding conservative accounts of perception. I then went on in the remainder of that chapter to argue that there are good reasons for thinking that testimonial conservatives are indeed committed to perceptual conservatism. To the extent that one finds conservatism about perception to be problematic—as arguments from earlier chapters demonstrate and as testimonial conservatives themselves recognise—one ought to find testimonial conservatism problematic also.

References

Bergmann, M. (2006). *Justification Without Awareness*. Oxford: Oxford University Press.

Brueckner, A. (2004). 'Strategies for Refuting Closure for Knowledge,' *Analysis*, 64: 333 - 335.

Burge, T. (1995). 'Content Preservation', *Philosophical Issues*, 6: 271 - 300.

Burge, T. (1996). 'Our Entitlement to Self-Knowledge', *Proceedings of the Aristotelian Society*, 96: 91 - 116.

Coady, C. A. J. (1992). *Testimony: A Philosophical Study*, Oxford: Clarendon Press.

Cohen, S. (1999). 'Contextualism, Skepticism and the Structure of Reasons', *Philosophical Perspectives*, 13: 57 - 89.

Cohen, S. (2002). 'Basic Knowledge and the Problem of Easy Knowledge', *Philosophy and Phenomenological Research,* 65: 309 - 329.

Coliva, A. (2012). 'Moderatism, Transmission Failures, Closure, and Humean Scepticism', in Dodd, D. & Zardini, E. (eds.), *Scepticism and Perceptual Justification*, Oxford: Oxford University Press, , 248 - 271.

Coliva, A. (2015). *Extended Rationality: A Hinge Epistemology,* Basingstoke: Palgrave-Macmillan.

Coliva, A. (2016). 'Which Hinge Epistemology?', *International Journal for the Study of Scepticism*, 6: 79 - 96.

Conee, E. & Feldman R. (1998). 'The Generality Problem for Reliabilism', *Philosophical Studies* 89: 1 - 29.

Davies, M. (2004). ‘Epistemic Entitlement, Warrant Transmission and Easy Knowledge’, *Proceedings of the Aristotelian Society*, *Supplementary* 78: 213 - 245.

Dodd, D. (2016). ‘Belief and Certainty’, *Synthese*, 194: 4597 - 4621.

Dretske, F. (1970). ‘Epistemic Operators’, *The Journal of Philosophy*. 67: 1007 - 1023.

Dretske, F. (2005). ‘The Case Against Closure,’ in Steup, M. & Sosa, E. (eds.), *Contemporary Debates in Epistemology*, Malden, Ma: Blackwell: 13 - 25.

Elstein, D. and Jenkins, C. (2019). ‘The Truth Fairy and the Indirect Epistemic Consequentialist’, in Graham, P. & Pedersen, N. (eds.), *Epistemic Entitlement,* Oxford: Oxford University Press, 344 - 360.

Feldman, R. and Conee, E. (1985). ‘Evidentialism’, *Philosophical Studies*, 48: 15 - 35.

Foley, R. (1992). ‘The Epistemology of Belief and the Epistemology of Degrees of Belief’, *American Philosophical Quarterly*, 29: 111-124.

Fricker, E. (1987). ‘The Epistemology of Testimony’, *The Proceedings of the Aristotelian Society*, 61: 57 - 83.

Fricker, E. (1995). ‘Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony’, *Mind*, 104: 393 - 411.

Fricker, E. (2002). ‘Trusting Others in the Sciences: A Priori or Empirical Warrant?’, *Studies in the History of Philosophy of Science*, 33: 373 - 383.

Gilbert, D.T., (1991). ‘How Mental Systems Believe’, *The American Psychologist*, 46: 107–119.

Gilbert, D. T., et al. (1990). ‘Unbelieving the Unbelievable’, *Journal of Personality and Social Psychology*, 59: 601–613.

Gilbert, D. T., et al. (1993). ‘You Can't Not Believe Everything You Read’, *Journal of Personality and Social Psychology*, 65: 221–233.

Graham, P. (2000[a]). ‘The Reliability of Testimony’, *Philosophy and Phenomenological Research,* 61 :695—709.

Graham, P. (2006). ‘Liberal Fundamentalism and Its Rivals’, in Lackey, J. and Sosa, E. (eds.), *The Epistemology of Testimony*. Oxford: Oxford University Press, 93—115.

Graham, P. (2010). ‘Testimonial Entitlement and the Function of Comprehension’, in Pritchard, D., Millar, A. & Haddock, A (eds.), *Social Epistemology*. Oxford: Oxford University Press, 148—174.

Graham, P. (2011). ‘Perceptual Entitlement and Basic Beliefs’, *Philosophical Studies*, 153: 467—475.

Graham, P. (2018[a]). ‘Formulating Reductionism about Testimonial Warrant and the Challenge from Childhood Testimony’, *Synthese*, 195: 3013—3033

Graham, P. (2018[b]). ‘Sincerity and the Reliability of Testimony: Burge on the A Priori Basis of Testimonial Entitlement’, in Stokke, A. & Michaelson, E. (eds.), *Lying: Language, Knowledge, Ethics*, *Politics*. Oxford: Oxford University Press, 85—112.

Graham, P. (2019). ‘The Function of Assertion and Social Norms’, in Goldberg, S. (ed.), *The Oxford Handbook of Assertion*. New York: Oxford University Press, 727—748.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Hajek, A. (2008). ‘Arguments for, or Against, Probabilism?’ *The British Journal for the Philosophy of Science*, 59: 793 - 819.

Hawthorne, J. (2005). ‘The case for closure’, in Steup, M., Turri, J. & Sosa, E. (eds.), *Contemporary Debates in Epistemology*, Malden: Blackwell, 26 – 42.

Jenkins, C. (2007). ‘Entitlement and Rationality’, *Synthese*, 157: 25 - 45.

Jope, M. (2019). ‘Closure Credence and Rationality: A Problem for Non-Belief Hinge Epistemology’, *Synthese*, 1 - 11.

Jope, M. (forthcoming[a]). ‘Intuitive Closure, Transmission Failure and Doxastic Justification’, in Jope, M. & Pritchard, D. (eds.), *New Perspectives on Epistemic Closure*, London: Routledge.

Jope, M. (forthcoming[b]). ‘On the Alleged Instability of Externalist Anti-Skepticism’, *The Journal of Philosophy*.

Jope, M. (forthcoming[c]). ‘The Symmetry Problem for Testimonial Conservatism’, *Synthese*.

Kornblith, H. (2002). *Knowledge and its Place in Nature*. Oxford: Oxford University Press.

Lackey, J. (2006). ‘It Takes Two to Tango: Beyond Reductionism and Non-Reductionism in the Epistemology of Testimony’, in Lackey, J. and Sosa, E. (eds.), *The Epistemology of Testimony*, Oxford: Oxford University Press, 160 - 189.

Lindsay et al. (2004). ‘True Photographs and False Memories’, *Psychological Science* 15: 149 - 154.

Lockard, M. (2016). ‘Closure Provides No Relief from the Problem of Easy Knowledge’, *Erkenntnis*, 79: 461 - 469.

Loftus et al. (1995). "The Formation of False Memories", *Psychiatric Annals*, 25: 720 - 725.

Luper, S. (2016). ‘Epistemic Closure’, *The Stanford Encyclopedia of Philosophy. Zalta, E. (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/closure-epistemic/>.*
McGlynn, A. (2013). [Review of the book ‘Mind Meaning and Knowledge by Annalisa Coliva], *Notre Dame Philosophical Reviews*, retrieved from https://ndpr.nd.edu/news/mind-meaning-and-knowledge-themes-from-the-philosophy-of-crispin-wright/

McGlynn, A. (2014). ‘On Epistemic Alchemy’, in Dodd, D. & Zardini, E. (eds.), *Scepticism and Perceptual Justification*, Oxford: Oxford University Press, 173-189.

McGlynn, A. (2016). ‘Epistemic Entitlement and the Leaching Problem.’ *Episteme*, 14: 89 - 102.

McKenna, R. (2015). ‘Contextualism in Epistemology’, *Analysis*, 75: 489 - 503.

Moore, G. E. (1939). ‘Proof of an External World’, *Proceedings of the British Academy*. 25, 273 - 300.

Moretti, L. & Piazza, T. (2018). ‘Transmission and Transmission Failure in Epistemology.’ *The Stanford Encyclopedia of Philosophy*. Zalta, E. (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/transmission-justification-warrant/>.

Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Plantinga, A. (1993). *Warrant and Proper Function*. New York: Oxford University Press.

Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.

Pritchard, D. (2007a). ‘Wittgenstein’s *On Certainty* and Contemporary Anti-Scepticism’, in Moyal-Sharrock, D. and Brenner, W. H. (eds.), *Readings of Wittgenstein’s On Certainty*, Basingstoke: Palgrave Macmillan, 189 - 224.

Pritchard, D. (2007b). ‘Anti-Luck Epistemology’, *Synthese*. 158: 277 - 298.

Pritchard, D. (2010). ‘Relevant Alternatives, Perceptual Knowledge and Discrimination’, *Nous*. 44 (2): 245 - 268.

Pritchard, D. (2012). *Epistemological Disjunctivism*. Oxford: Oxford University Press.

Pritchard, D. (2016a). *Epistemic Angst*. Princeton: Princeton University Press.

Pritchard, D. (2016b). ‘Epistemic Angst’, *Philosophical and Phenomenological Research*, 93: 70 - 90.

Pritchard, D. and Ranilli, C. (2016). ‘On Metaepistemological Scepticism’, in Bergmann, M. & Coppenger, B. (eds.), *Intellectual Assurance: Essays on Traditional Epistemic Internalism*, Oxford: Oxford University Press, 205 - 223

Pryor, J. (2000). ‘The Skeptic and the Dogmatist’, *Noûs*, 34: 517 - 549.

Pryor, J. (2004). ‘What’s Wrong with Moore’s Argument’, *Philosophical Issues*. 14: 349-378.

Pryor, J. (2012). ‘When Warrant Transmits’ in Dodd, D. & Zardini, E. (eds.), *Scepticism and Perceptual Justification*, Oxford: Oxford University Press, 269-303.

Rabinowitz, D. (2011). ‘The Safety Condition for Knowledge.’ *The Internet Encyclopaedia of Philosophy*. Retrieved from https://www.iep.utm.edu/safety-c/

Reichenbach, H. (1938). *Experience and Prediction*. Chicago: Chicago University Press.

Russell, B. (2005 [1921]). *The Analysis of Mind*. London: Routledge.

Rysiew, P. (2012). ‘Beyond Worlds: Communication, Truthfulness and Understanding”, Episteme 4: 285—304.

Simion, M. & Kelp, C. (2018). ‘How to Be an Anti-Reductionist’, *Synthese,* 197: 2849 - 2866.

Silins, N. (2005). ‘Transmission Failure Failure’, *Philosophical Studies,* 126: 71 - 102.

Smith, M. (2009). ‘Transmission Failure Explained’, *Philosophy and Phenomenological Research,* 79: 164 - 189.

Smith, M. (2016). ‘Scepticism by a Thousand Cuts’, *International Journal for the Study of Scepticism*, 6: 44 - 52.

Smith, M. (2019), ‘Full-Blooded Entitlement’, in Graham, P. and Pedersen, N. (eds.), *New Essays on Entitlement*, Oxford: Oxford University Press, 281 - 296.

Sosa, E. (1994). ‘Philosophical Scepticism and Epistemic Circularity’, *Proceedings of the Aristotelian Society*, *Supplementary Volumes*, 68: 263 - 290.

Sosa, E. (1999a). ‘How to Defeat Opposition to Moore’, *Philosophical Perspectives,* 13: 137 - 149.

Sosa, E. (1999b). ‘How Must Knowledge be Modally Related to What is Known?’ *Philosophical Topics*, 373 - 384.

Stroud, B. (1994). ‘Scepticism, ‘Externalism’, and the Goal of Epistemology,’ *Proceedings of the Aristotelian Society*, *Supplementary Volumes*, 68: 291 - 307.

Turri, J. (2010). ‘On the Relationship between Propositional and Doxastic Justification.’ *Philosophical and Phenomenological Research*, 80: 312 - 326.

Wade et al. (2001). ‘A Picture is Worth a Thousand Lies: Using False Photographs to Create False Childhood Memories. *Psychonomic Bulletin & Review*. 9: 597 - 603.

Warfield, T. (2004). ‘When Epistemic Closure Does and Does Not Fail: A Lesson From the History of Philosophy,’ *Analysis*, 64: 41 - 43.

Weatherson, B. (2019). *Normative Externalism*. Oxford: Oxford University Press.

White, R. (2006). ‘Problems for Dogmatism’ *Philosophical Studies*, 131: 525 - 557.

Wittgenstein, L. (1969) *On Certainty*, Anscombe, G. E. M. & Wright, G. H. (eds.). Oxford: Blackwell.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Williamson, T. (2009). 'Reply to Alvin Goldman.' In Greenough, P. & Pritchard, D. (eds.), *Williamson on Knowledge*, Oxford: Oxford University Press, 305 - 312.

Wright, C. (1986). 'Facts and Certainty', in *Proceedings of the British Academy*, 71: 429 - 472.

Wright, C. (1991). 'Scepticism and Dreaming: Imploding the Demon.' *Mind*, 100: 87 - 116.

Wright, C. (2000). 'Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Moore's Proof.' *Philosophical Issues*, 10: 140 - 163.

Wright, C. (2002). '(Anti-)Sceptics Simple and Subtle: G. E. Moore and John McDowell.' *Philosophy and Phenomenological Research*, 65: 330 - 348.

Wright, C. (2003), 'Some Reflections on the Acquisition of Warrant by Inference,' in Nuccetelli, S. (ed.), *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge: MIT Press, 57 - 78.

Wright, C. (2004). 'On Epistemic Entitlement: Warrant for Nothing (and Foundations for Free?)', *Proceedings of the Aristotelian Society*, *Supplementary* 78: 167 - 212.

Wright, C. (2008). 'Internal-External: Doxastic Norms and the Defusing of Skeptical Paradox', *The Journal of Philosophy*, 105: 501 - 517.

Wright, C. (2012). 'Replies', in Coliva, A. (ed.), *Mind, Meaning and Knowledge; Themes from the Philosophy of Crispin Wright*, Oxford: Oxford University Press, 451 - 486.

Wright, C. (2014) 'On Epistemic Entitlement (II): Welfare State Epistemology', in Dodd, D. and Zardini, E. (eds.), *Scepticism and Perceptual Justification*, New York: Oxford University Press, 213 - 247.